

THÈSE

présentée

**A LA FACULTÉ DES SCIENCES ET DES TECHNIQUES
DE L'UNIVERSITÉ DE BESANÇON**

pour obtenir le

DIPLOME DE DOCTEUR INGÉNIEUR

par

Michel FILLEAU

Ingénieur ENSCMB

**DISPOSITIF D'ANALYSE EN TEMPS RÉEL
DE LA MÉLODIE DE LA PAROLE**

Soutenue le 23 septembre 1975 devant la Commission d'Examen :

MM. F. LHOTE

Président

J.M. KAUFFMANN
J.C. LAFON
M. ROSSI
M. WAJSKOP

} Examineurs

THÈSE

présentée

**A LA FACULTÉ DES SCIENCES ET DES TECHNIQUES
DE L'UNIVERSITÉ DE BESANÇON**

pour obtenir le

DIPLOME DE DOCTEUR INGÉNIEUR

par

Michel FILLEAU

Ingénieur ENSCMB

**DISPOSITIF D'ANALYSE EN TEMPS RÉEL
DE LA MÉLODIE DE LA PAROLE**

Soutenue le 23 septembre 1975 devant la Commission d'Examen :

MM. F. LHOTE

Président

J.M. KAUFFMANN
J.C. LAFON
M. ROSSI
M. WAJSKOP

} Examineurs



UNIVERSITE DE BESANCON

Président : Monsieur THIEBAUT Jean

FACULTE DES SCIENCES ET DES TECHNIQUES

-:-:-:-:-

DOYEN

Mr BERNARD Jean : Chimie-Physique
Professeur :

ASSESSEUR

Mr PARIZET Jean : Mathématiques
Professeur :

PROFESSEURS

Mrs CHATELET François : Mathématiques
MESNAGE Pierre : Chronométrie
PERROT Roger : Chimie Générale
PLUVINAGE Philippe : Physique Théorique
QUANTIN André : Pédologie
CHALEAT Raymond : Mécanique Appliquée
RIPPLINGER Jean : Physiologie animale
GREMILLARD Jean : Mathématiques
VIENOT Jean-Charles : Physique Gle-Optique
THIEBAUT Jean : Pétrographie-Minéralogie
GOMOT Lucien : Zoologie et Embryologie
BENNETON Gaston : Mathématiques
GALATRY Louis : Physique Moléculaire
THEOBALD Jean-Gérard : Spectroscopie Hertz.
CERUTTI Ernest : Chimie Appliquée
DEVIN Claude : Chimie 1er cycle
CAPODANNO Pierre : Mécanique Théorique
BANTEGNIE Robert : Mathématiques
ROBERT Daniel : Physique Moléculaire
ROBERT Jacques : Mathématiques I
CHAUVE Pierre : Géologie-Minéralogie
RANGHEARD Yves : Géologie

PROFESSEURS SANS CHAIRE

Mrs POTIER Robert : Physique des Solides
REAL Pierre : Ecologie animale
GOUARNE René : Mathématiques
MERIGOUX Henri : Cristallographie et
Synthèses minérales
OLIVIER Marcel : Electronique Quantique
MONTAGNER Hubert : Psychophysiologie
BULABOIS Jean : Physique Gle-Optique
GAUDEMER Yves : Biochimie
LAUDE Bernard : Chimie Organique
MIELLOU Jean-Claude : Analyse Numérique
ROBERT Guy : Electrochimie

DOYENS HONORAIRES

Mrs GLANGEAUD, JACQUEMAIN,
CHATELET, THIEBAUT

PROFESSEURS HONORAIRES

Mrs BAILLAUD, DUFFIEUX, THIRY,
TRILLAT, UEBERSFELD,
VINCENSINI, DREYFUSS
TRONCHET, THEOBALD N.

MAITRE DE CONFERENCES HONORAIRE

Mr GRANIER

MAITRES DE CONFERENCES

Mrs BIDAULT Michel : Taxo.Exp.
DUBOUCHET Jacques : Botanique
MARLE Charles : Mathém.
WEIL Michel : Mathém.
HAYLI Avram : Astrono.
BROQUET Paul : Géologie
GRAS Georges : Mathém.
TREHEL Michel : Informa.

CHARGE D'ENSEIGNEMENT

Mr MILLET Bernard

RESPONSABLE DES SERVICES ADMINISTRATIFS

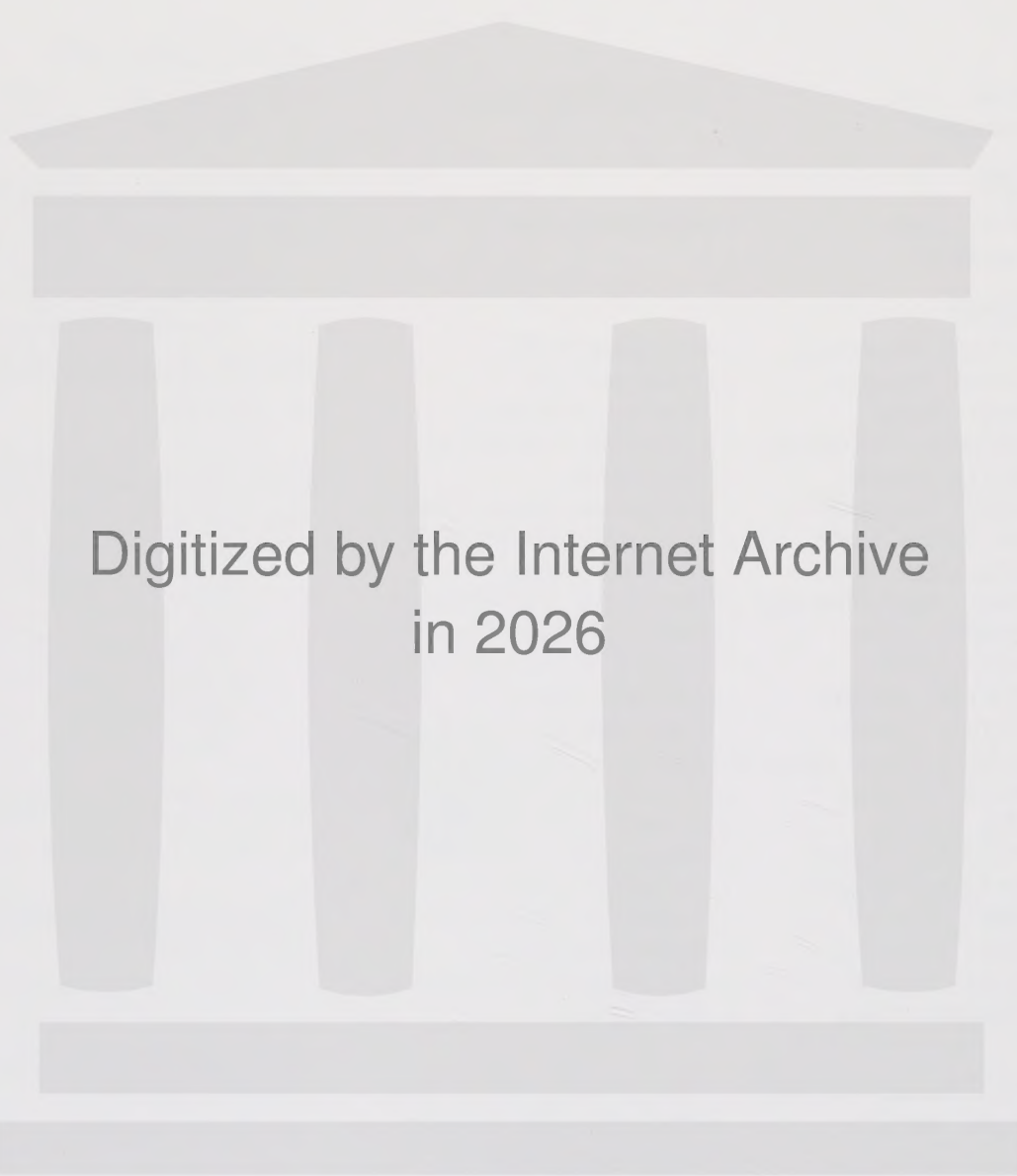
Mr RUNGE Jean-Marie

PROFESSEUR SANS CHAIRE DE L'E.N.S.C.M.B.

Mr LHOTE François

MAITRE DE CONFERENCES

Mr BESSON Raymond



Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41546611_0022

S O M M A I R E

	<u>Pages</u>
<u>INTRODUCTION</u>	
 <u>CHAPITRE 1 : LA PHONATION ET LES VARIATIONS DE HAUTEUR DANS LA PAROLE</u>	 1
<u>1.1. L'APPAREIL PHONATOIRE</u>	1
1.1.1. <u>Description</u>	1
1.1.2. <u>Le larynx</u>	3
1.1.2.1. <u>Anatomie</u>	3
1.1.2.2. <u>Genèse du mouvement des cordes vocales</u>	3
1.1.2.3. <u>Rôle du larynx dans la production de la parole</u>	4
1.1.3. <u>Fonctionnement de l'appareil phonatoire</u>	5
 <u>1.2. STRUCTURE ACOUSTIQUE ET LINGUISTIQUE DE LA PAROLE</u> ..	 5
1.2.1. <u>Analyse de la parole</u>	5
1.2.2. <u>Paramètres linguistiques</u>	6
1.2.2.1. <u>Le phonème</u>	6
1.2.2.2. <u>Les paramètres prosodiques</u>	7

1.3. INFLUENCE LINGUISTIQUE DES VARIATIONS DE HAUTEUR	
<u>DE LA PAROLE</u>	8
1.3.1. <u>Au niveau de la phrase : mélodie de la voix</u> <u>et intonation</u>	8
1.3.2. <u>Au niveau du phonème</u>	9
1.4. <u>LES ANALYSEURS DE MELODIE</u>	9
1.4.1. <u>Spécifications générales concernant les analyseurs</u> <u>de mélodie</u>	10
1.4.2. <u>Les différents types d'analyseurs existants</u> .	11
1.4.2.1. <i>Analyse dans le domaine fréquentiel.</i>	12
1.4.2.2. <i>Analyse dans le domaine temporel</i> ...	14
1.4.3. <u>Un nouveau type d'analyseur</u>	16
 CHAPITRE 2 : <i>DETECTION ELECTROGLOTTOGRAPHIQUE DU MOUVEMENT</i> <i>DES CORDES VOCALES</i>	18
2.1. <u>LES METHODES D'OBSERVATION DU MOUVEMENT DES CORDES</u> <u>VOCALES</u>	18
2.1.1. <u>Méthodes d'observation directe</u>	18
2.1.2. <u>Méthodes d'observation indirecte</u>	19
2.1.3. <u>L'électroglottographe de Fabre</u>	20
2.2. <u>CARACTERISTIQUES DE L'IMPEDANCE ELECTRIQUE EQUIVA-</u> <u>LENTE AU LARYNX</u>	22
2.2.1. <u>Comportement des tissus vivants en courant</u> <u>alternatif</u>	24
2.2.2. <u>Schéma équivalent au larynx</u>	24

2.2.3. Etude des variations de l'impédance du larynx au cours de la phonation	25
2.2.4. Conclusion	29
<u>2.3. INTERPRETATION DU GLOTTOGRAMME</u>	29
<u>2.4. ETUDE ET REALISATION D'UN ELECTROGLOTTOGRAPHE</u>	31
2.4.1. Dispositif de modulation en amplitude de la tension porteuse	32
2.4.1.1. Principe et montages possibles ...	32
2.4.1.2. Etude du montage adopté	35
2.4.2. Détection de la modulation d'amplitude	39
2.4.2.1. But et principe de la détection ..	39
2.4.2.2. Détection des signaux forts à l'aide d'une diode	42
2.4.2.3. Circuits détecteurs à transistors .	43
2.4.3. Etude du filtrage du glottogramme	47
2.4.4. Notice technique de l'électroglottographe .	52
2.4.4.1. Détails de construction	52
2.4.4.2. Bornes d'accès et commandes	55
2.4.4.3. Méthode d'utilisation	56
 <u>CHAPITRE 3 : DISPOSITIF DE MESURE DE FREQUENCE INSTANTANEE .</u>	59
 <u>3.1. MISE EN FORME DU GLOTTOGRAMME</u>	59
<u>3.2. CARACTERISTIQUES IMPOSEES A L'APPAREIL DE MESURE .</u>	64
<u>3.3. LES METHODES DE MESURE DE FREQUENCE INSTANTANEE ..</u>	65

3.3.1. Méthodes analogiques	65
3.3.2. Méthodes numériques	67
 3.4. <u>DISPOSITIF DIGITAL DE MESURE LOGARITHMIQUE DE</u> <u>FREQUENCE INSTANTANEE</u>	 68
3.4.1. Principe général	68
3.4.2. Précision de la mesure	70
3.4.3. Générateur d'impulsions à distribution ex- ponentielle	73
3.4.3.1. Synthèse globale absolue	73
3.4.3.2. Synthèse absolue par octave	75
3.4.3.3. Synthèse incrémentale par octave .	77
3.4.4. Réalisation de l'appareil	80
3.4.4.1. Horloge	80
3.4.4.2. Générateur d'impulsions à distri- bution exponentielle	83
3.4.4.3. Compteur d'intervalles élémentaires	87
3.4.4.4. Registre de sortie et convertisseur digital-analogique	88
3.4.4.5. Dispositif d'initialisation	92
3.4.4.6. Conclusion	92
 CHAPITRE 4 : <u>APPLICATIONS GENERALES DE LA CHAINE D'ANALYSE</u> <u>DE MELODIE</u>	 95
4.1. <u>LA COURBE DE MELODIE</u>	95
4.1.1. L'analyseur de mélodie	95
4.1.2. Présentation des mesures	96

	<u>Pages</u>
4.2. <u>ANALYSE DES VARIATIONS DE HAUTEUR DE LA PAROLE ..</u>	98
4.2.1. <u>Introduction</u>	98
4.2.2. <u>Détermination de schémas intonatifs simples ..</u>	100
4.2.3. <u>Micromélorie</u>	101
4.3. <u>SYNTHESE DE LA MELODIE DE LA PAROLE</u>	102
4.3.1. <u>L'analyse-synthèse de la parole par vocoder ..</u>	102
4.3.2. <u>Détection et reproduction de la mélodie ..</u>	104
4.3.3. <u>Autres applications du générateur de mélodie ..</u>	106
4.3.3.1. <i>Générateur des notes de la gamme musicale</i>	106
4.3.3.2. <i>Générateur de signaux à variation de fréquence logarithmique</i>	107
4.4. <u>CONCLUSION</u>	108
 <u>CHAPITRE 5 : L'ANALYSE DES LANGUES A TONS</u>	 109
5.1. <u>LES LANGUES A TONS</u>	110
5.2. <u>ETUDE DES TONS DU YORUBA</u>	112
5.3. <u>DISPOSITIF D'ENTRAINEMENT A LA PRODUCTION ET A LA PERCEPTION DES VARIATIONS TONALES DE LA VOIX</u>	11
 <u>ANNEXE I : REALISATION D'UN SEGMENTATEUR</u>	 119
 <u>CONCLUSION</u>	
 <u>BIBLIOGRAPHIE</u>	

I N T R O D U C T I O N

Les linguistes ont longtemps négligé la contribution des paramètres prosodiques dans l'information véhiculée par la parole. Mais leurs recherches tendent actuellement à montrer que ces faits, et surtout l'intonation permettent d'expliquer des problèmes linguistiques jusque là insolubles.

Une des causes principales du démarrage tardif de l'étude systématique de la prosodie a été l'absence d'appareillages expérimentaux adaptés.

La détermination de la variation de hauteur dans la parole, qui constitue un paramètre important de la communication humaine, n'a pu être réalisée pendant plus d'un demi-siècle que par le dépouillement manuel d'enregistrements graphiques obtenus au kymographe (inventé par Rousselot en 1908) ou par l'analyse auditive directe.

L'apparition récente de techniques qui permettent d'appliquer au signal de parole les notions d'analyse spectrale et d'analyse par autocorrélation a conduit à la réalisation de nombreux appareils dont les principes s'inspirent de l'une, de l'autre ou d'une combinaison des deux théories. Ces analyseurs sont cependant restés pour la plupart à l'état de prototypes, car les plus simples ont des performances médiocres, et les plus élaborés procèdent à de longs calculs nécessitant un ordinateur. De plus, il en est très peu qui fonctionnent en temps réel en conservant de bonnes conditions de fiabilité et de précision.

L'étude que nous proposons, réalisée au Laboratoire d'Automatique de l'Ecole Nationale Supérieure de Chronométrie et de Micromécanique sur une proposition du Laboratoire de Phonétique de la Faculté des Lettres de Besançon concerne un système simple, fiable et précis de détermination en temps réel de la hauteur des faits laryngés au cours de la parole utilisant la technique d'électroglottographie introduite par Ph. Fabre (1957).

Un premier chapitre rappelle succinctement le fonctionnement de l'appareil phonatoire en insistant sur la place et l'importance qu'occupe la source laryngée, et situe les problèmes d'analyse des variations de hauteur tant au niveau de la phrase (mélodie, intonation) qu'au niveau du phonème (micromélodie, commande d'un vocoder par une source vocale, langues à tons).

Les deux chapitres suivants rapportent l'étude et la réalisation de chaque module de l'analyseur de mélodie que nous avons réalisé.

Le quatrième chapitre montre, sur des résultats d'expériences, les possibilités d'utilisation de l'appareil : analyse de l'intonation d'une phrase, mesure de la hauteur des phonèmes.

Le dernier chapitre est consacré à l'adaptation du système à l'étude des langues à tons. Il se termine sur l'application du dispositif expérimental à l'apprentissage des variations tonales de la parole par visualisation oscilloscopique immédiate de la "courbe mélodique". Cette méthode est particulièrement intéressante pour faciliter la perception et la production de voyelles des langues à tons.

C H A P I T R E 1

LA PHONATION ET LES VARIATIONS DE HAUTEUR DANS LA PAROLE

Le début de ce chapitre, par un bref rappel de la constitution et du fonctionnement de l'appareil phonatoire, permet de mettre en valeur la nécessité de l'étude des variations de hauteur dans la parole.

Le mot hauteur désigne, usuellement, la sensation auditive due à la perception de la fréquence d'un son périodique. Mais l'expérience montre que la hauteur est encore perçue lorsque le signal périodique ne dure qu'un court instant. C'est le cas des sons voisés de la parole. La sensation de hauteur à l'audition d'un message parlé est donc liée directement à la durée qui sépare les impulsions laryngées successives.

1.1. L'APPAREIL PHONATOIRE

1.1.1. Description

Les organes de la production des sons de la parole, dont le schéma simplifié est donné à la figure 1, sont constitués d'un conduit divisé en trois cavités (le pharynx, la cavité buccale, la cavité nasale) parcourues par un courant d'air provenant de l'expulsion du contenu des poumons à travers la trachée artère et le larynx. La partie la plus importante du larynx est constituée par les cordes vocales.

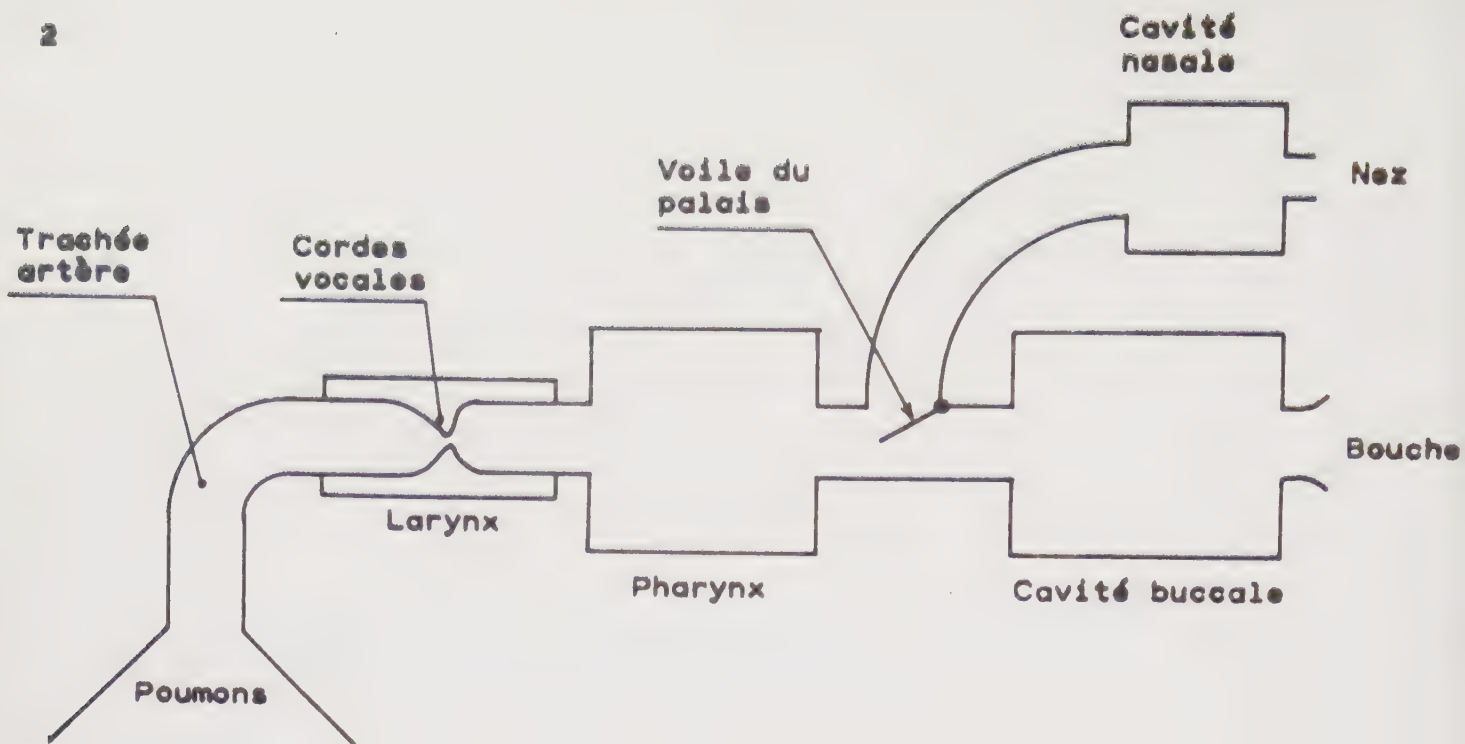


Figure 1 - SCHEMA DE L'APPAREIL PHONATOIRE .

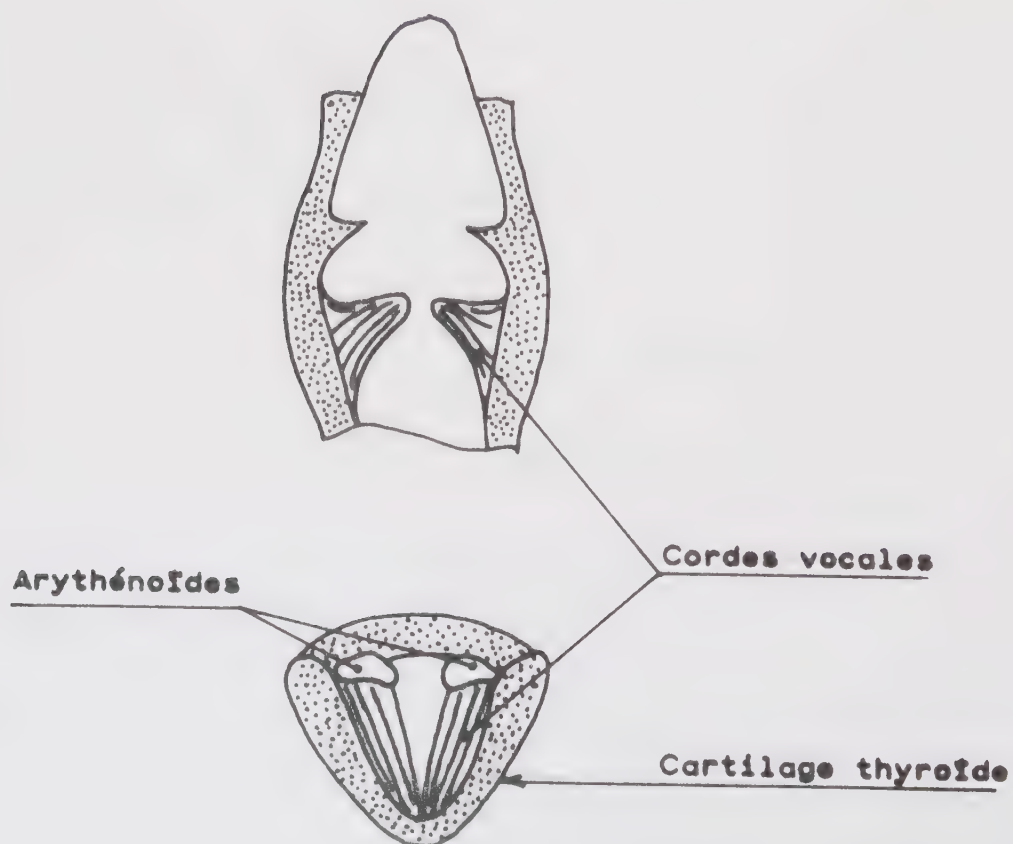


Figure 2 - ANATOMIE DU LARYNX .

Le locuteur peut agir à sa guise sur la forme et le volume du pharynx, par déformation des muscles qui l'entourent et de la cavité buccale, par la position de la langue, des lèvres et du voile du palais, lequel commande également le couplage entre les cavités buccale et nasale.

1.1.2. Le larynx

Le larynx, source des sons "voisés", est l'organe essentiel de la production des variations de hauteur dans la parole.

1.1.2.1. *Anatomie*

Le larynx est un renflement d'anneaux cartilagineux situé à l'extrémité de la trachée artère ; il contient deux membranes musculaires épaisses, les cordes vocales, pouvant s'accoler et s'écarter transversalement ; ces dernières, attachées en un point commun à l'avant du larynx ont chacune une extrémité liée à un petit cartilage mobile (aryténoïde) (figure 2). Les mouvements possibles des cordes vocales déterminent ainsi une ouverture triangulaire de section variable, placée sur le trajet de l'air venant des poumons et appelée glotte.

1.1.2.2. *Genèse du mouvement des cordes vocales*

Si, pendant un temps, la théorie neurochronaxique de R. Husson et quelques autres propositions ont divisé les physiologistes qui cherchent à expliquer le fonctionnement du larynx, il semble actuellement admis que la mise en mouvement des cordes vocales réponde en grande partie à la théorie myo-élastique.

Selon cette théorie, la source laryngée se comporte comme un oscillateur à relaxation. La glotte, normalement ouverte pendant la respiration ordinaire, est placée en position de phonation (quasiment fermée) par l'action d'un influx nerveux sur un muscle approprié. La pression sous-glottique de l'air expulsé des poumons augmente jusqu'à un seuil où se produit l'écartement brusque des cordes vocales ; l'air passe et la pression diminue, ce qui permet à la glotte de se refermer et ainsi de suite. La durée d'un cycle est fixée par la pression de l'air, la masse et la tension des cordes vocales.

1.1.2.3. *Rôle du larynx dans la production de la parole*

L'air provenant des poumons traverse l'orifice de section variable constitué par la glotte. Lorsque les cordes vocales sont animées d'un mouvement rythmé, le courant d'air est haché en une suite de bouffées appelées impulsions laryngées. Cette suite d'impulsions, transformée par les résonances des cavités sus-glottiques en un signal de parole, constitue la variable d'excitation du conduit vocal pour un grand nombre de sons vocaux (les voyelles et les consonnes voisées).

La sensation de hauteur ressentie à l'audition d'un message parlé est identique, à un instant donné, à celle que produirait un son périodique de fréquence (f) égale à l'inverse de l'intervalle de temps (T) nécessaire à l'exécution d'un cycle glottique.

C'est cette constatation qui fait habituellement nommer la quantité $f = 1/T$ fréquence fondamentale de la voix. Cette expression ne peut,

cependant, s'appliquer qu'à des signaux périodiques, ce qui n'est pas le cas des suites d'impulsions laryngées. C'est la raison pour laquelle nous préférons utiliser les termes de "fréquence mélodique instantanée".

1.1.3. Fonctionnement de l'appareil phonatoire

Le fonctionnement de l'appareil phonatoire se fait suivant trois modes principaux : voisement, constriction, occlusion.

Nous venons de voir que la production des voyelles nécessite l'excitation des cavités résonantes du conduit vocal par des suites d'impulsions laryngées ; c'est le mode "voisé" ; il est à l'origine des sensations de hauteur dans la parole.

Les consonnes constrictives [ʃ, f, s, etc...] sont engendrées par des sources de bruit dues à des turbulences de l'air en des endroits du conduit vocal formant des constriction. Certaines de ces consonnes [v, j, z] s'accompagnent de voisement.

Les consonnes occlusives [t, k, p, etc...] proviennent de l'expulsion brutale en un point déterminé du conduit d'un volume d'air préalablement comprimé. Comme ci-dessus, certaines de ces consonnes [b, d, g, m, n] s'accompagnent de voisement.

Le couplage de la cavité nasale à l'aide du voile du palais permet d'obtenir les voyelles et les consonnes nasalisées [ã, ɔ̃, ɛ̃, œ̃, m̃, ñ]

1.2. STRUCTURE ACOUSTIQUE ET LINGUISTIQUE DE LA PAROLE

1.2.1. Analyse de la parole

Les recherches concernant les mécanismes de la phonation et la connaissance des langues reposent sur le développement de l'analyse de la parole. Celle-ci consiste à déterminer un nombre minimum de paramètres qui permettent de décrire au plus juste les contenus acoustique et linguistique du message parlé, ainsi que leurs rapports réciproques.

Les acousticiens, par l'étude de la perception des sons de la parole, définissent le signal à l'aide de quatre types de variables acoustiques : la hauteur, l'intensité, la durée et le timbre des sons. Les trois premiers correspondent aux sensations auditives dues, respectivement, aux paramètres de fréquence, d'amplitude et de longueur temporelle du signal ; le timbre est caractérisé par le contenu spectral et les transitoires du son considéré.

Les phonéticiens, pour leur part, ont mis en évidence deux notions linguistiques : les phonèmes et les faits prosodiques qui constituent la forme de l'information linguistique contenue dans le message parlé. On les caractérise physiquement à l'aide des variables acoustiques, articulatoires ou perceptuelles.

1.2.2. Paramètres linguistiques

1.2.2.1. *Le phonème*

Le phonème représente l'information linguistique véhiculée par une sensation sonore qui donne à l'oreille une impression globale d'unité de perception.

Du point de vue acoustique, le phonème est défini par son intensité, sa hauteur, sa durée et son contenu spectral. Celui-ci permet de classer les phonèmes en deux groupes :

- les phonèmes vocaliques, dont la production s'accompagne de vibrations des cordes vocales; leur spectre d'amplitude contient des raies harmoniques dominant le bruit et fait apparaître des zones de fréquences où l'amplitude est renforcée, appelées "formants".
- les phonèmes non vocaliques, pour lesquels les cordes vocales sont immobiles ; leur spectre est essentiellement formé de bruit.

Les phonèmes constituent, pour une langue donnée, un véritable code du langage parlé.

1.2.2.2. *Les paramètres prosodiques*

Les paramètres prosodiques sont principalement: l'accent, le rythme, l'intonation et les tons.

L'information qu'ils véhiculent concerne l'agencement acoustique global de la phrase, par opposition au phonème dont l'apport est très local.

L'accent est un paramètre contrastif. Il permet de mettre certaines syllabes en valeur par rapport aux autres ; il contribue ainsi au modelage du sens de la phrase. Sur le plan acoustique, il est réalisé, suivant les langues, par la variation de l'une ou plusieurs des trois variables acoustiques (durée, intensité, hauteur).

Le rythme caractérise la distribution des accents et des pauses. Il permet de délimiter clairement des groupes de mots significatifs et intervient dans le naturel de la voix.

L'intonation est le paramètre prosodique le plus important. Les phonéticiens considèrent qu'elle contribue fortement à l'apport d'information dans le message parlé, mais n'ont pas encore réussi à se mettre d'accord sur sa fonction linguistique. Du point de vue acoustique, l'intonation dépend des variations de hauteur, de durée et d'intensité des sons dans la phrase.

1.3. INCIDENCE LINGUISTIQUE DES VARIATIONS DE HAUTEUR DE LA PAROLE

Après avoir déterminé l'origine des variations de hauteur de la voix par l'examen des organes de la phonation et suggéré brièvement la nature acoustique et linguistique de la parole, nous allons préciser l'action de la hauteur sur l'information du message parlé à différents niveaux qui motiveront l'évolution de l'étude de l'appareillage dont ce mémoire est l'objet.

1.3.1. Au niveau de la phrase : mélodie de la voix et intonation

Les phonéticiens distinguent, dans les productions phonatoires articulées, "la voix" qui englobe l'ensemble du signal acoustique rayonné par la bouche et la "parole" qui en est le contenu informationnel. La voix est un signal personnalisé ; elle résulte de l'adjonction à la parole de traits propres au locuteur.

Nous appellerons "mélodie" les variations de hauteur de la voix, alors que celles de la parole constitueront "l'intonation". L'une et l'autre sont superposées dans la phrase qu'entend l'oreille. L'interprétation d'une analyse expérimentale des variations de hauteur de la voix qui est seule accessible physiquement, en vue de déterminer le message intonatif, doit donc être faite avec prudence.

1.3.2. Au niveau du phonème

La hauteur d'un phonème vocalique est définie par la perception auditive de la composante fondamentale (ou d'ordre 1) de son spectre d'amplitude qui correspond au son laryngé.

Un même phonème peut être accompagné de sensations de hauteur différentes suivant sa place dans le mot et dans la phrase. Il est intéressant, du point de vue linguistique, d'étudier la distribution relative des hauteurs d'un phonème en fonction de ceux qui l'entourent et du sens de la phrase.

Dans le domaine de l'analyse-synthèse de la parole par vocoder, la détermination automatique de la hauteur des phonèmes peut conduire à une amélioration sensible de la qualité du signal obtenu par une commande dynamique de la source vocale.

Au début et à la fin d'un phonème se produit une légère altération transitoire de la ligne générale des variations de hauteur de la parole appelée "micro-mélodie". Elle est due à des phénomènes physiologiques. Elle se manifeste de façon analogue dans toutes les langues et est indépendante de la mélodie du message parlé.

La micromélodie est un paramètre qui influe sur le naturel de la voix et n'est pas contrôlable par le locuteur.

Il convient également de citer ici le cas des langues à tons. Les variations de hauteur à l'intérieur même des voyelles de ces langues, constituent des traits oppositifs, porteurs d'information linguistique : un mot, dont la transcription phonétique est donnée, change de sens avec le schéma tonal des ses phonèmes vocaliques.

1.4. LES ANALYSEURS DE MELODIE

L'examen rapide de l'influence des variations de hauteur de la voix sur le message parlé montre la nécessité d'en effectuer l'analyse expérimentale. C'est la raison pour laquelle les efforts des chercheurs ont porté depuis une quinzaine d'années, sur la réalisation de systèmes permettant d'obtenir le schéma tonal de la voix en fonction du temps : les analyseurs de mélodie. Léon et Martin[31] en ont recensé quelques 150 construits dans le monde entier. Malheureusement, les principes utilisés nécessitent, pour l'obtention de performances raisonnables, un équipement complexe et coûteux, comprenant souvent un ordinateur, ce qui interdit l'acquisition de ces appareils à de nombreux petits laboratoires.

1.4.1. Spécifications générales concernant les analyseurs de mélodie

L'obtention du graphe exact des variations de hauteur de la voix en fonction du temps nécessite de la part de l'analyseur quatre caractéristiques fondamentales :

- précision de la détection du signal à mesurer et fiabilité
- précision de la mesure
- aptitude à la restitution correcte des variations brusques et de grande amplitude de la hauteur
- fonctionnement en temps réel sans préjudice pour les autres caractéristiques

Le domaine d'analyse doit couvrir toute la bande de "fréquences" pouvant être produites par les mouvements des cordes vocales. Les phonéticiens s'accordent pour fixer les limites inférieure et

supérieure de la hauteur des voix humaines dans la communication parlée normale à 80 et 600 Hz, les valeurs moyennes étant (d'après FANT - 1959) :

- 125 Hz pour les hommes
- 250 Hz pour les femmes
- 300 Hz pour les enfants

La plus petite variation de fréquence que puisse percevoir l'oreille est fixée par les musiciens à 1 comma (le comma est un rapport de fréquences de $81/80 = 1,0125$) soit une différence de 1,25 % de la valeur nominale. Cette sensibilité est rarement atteinte par une oreille non exercée, et la détermination des variations de hauteur de la voix à 2 % est satisfaisante (sauf pour la micromélodie qui demande une analyse plus fine).

La variation des intervalles de temps de récurrence des impulsions laryngées, en dehors des instants de mise en mouvement et d'arrêt des cordes vocales peut atteindre 20 % (selon LEON et MARTIN) de la durée du cycle, d'une impulsion à l'autre.

1.4.2. Les différents types d'analyseurs existants

Les analyseurs construits jusqu'à ce jour utilisent tous les notions d'analyse spectrale du signal rayonné par la bouche pour le domaine fréquentiel et d'analyse par autocorrélation pour le domaine temporel. Les deux théories sont étroitement liées. Dans "Prolégomènes à l'étude des structures intonatives", dont s'inspirent les pages suivantes, LEON et MARTIN proposent une copieuse bibliographie se rapportant aux réalisations d'analyseurs de mélodie.

1.4.2.1. Analyse dans le domaine fréquentiel

L'analyse spectrale d'un signal $f(t)$ est basée sur la détermination de sa transformée de Fourier.

Considérons une fonction $f(t)$, définie en tout point de l'axe réel, continue ainsi que sa dérivée, sauf éventuellement en un nombre fini de discontinuités de première espèce ; l'intégrale de Fourier de cette fonction a un sens pour $f(t)$ et donne sa transformée de Fourier $F(\omega)$ appelée densité spectrale complexe ou plus simplement spectre de $f(t)$. Elle s'exprime par :

$$F(\omega) = \int_{-\infty}^{+\infty} f(t) e^{i\omega t} dt$$

Dans le cas où $f(t)$ présente un caractère périodique (ou "pseudo périodique" s'il est limité dans le temps), il apparaît, dans sa densité spectrale d'amplitude $|F(\omega)|$ des raies correspondant à des fréquences multiples les unes des autres, et dont le plus petit multiple est responsable de la sensation de hauteur à l'audition de $f(t)$.

Aucun dispositif physique ne permettant de calculer une intégrale à bornes infinies, on utilise le spectre quasi-instantané $H(\theta, \omega)$, prélevé autour d'un instant d'observation θ à l'aide d'une fonction de pondération $p(t)$ telle que :

$$p(t) = 0 \quad \text{si} \quad t \notin [-T, T]$$

$$p(t) \neq 0 \quad \text{si} \quad t \in [-T, T] \quad \text{et}$$

$$H(\theta, \omega) = \int_{\theta-T}^{\theta+T} f(t) p(t-\theta) e^{-i\omega t} dt$$

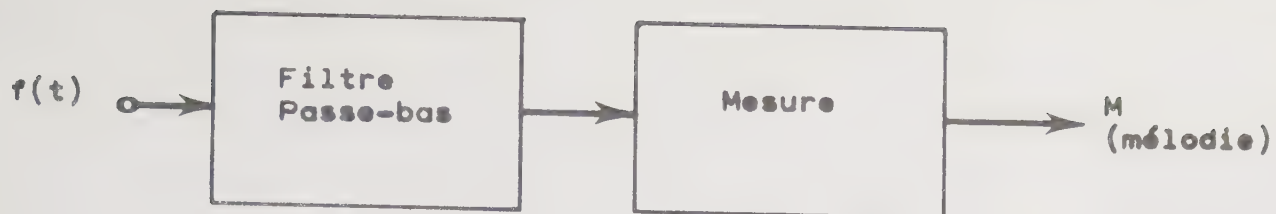


Figure 3 - ANALYSEUR A UN FILTRE .

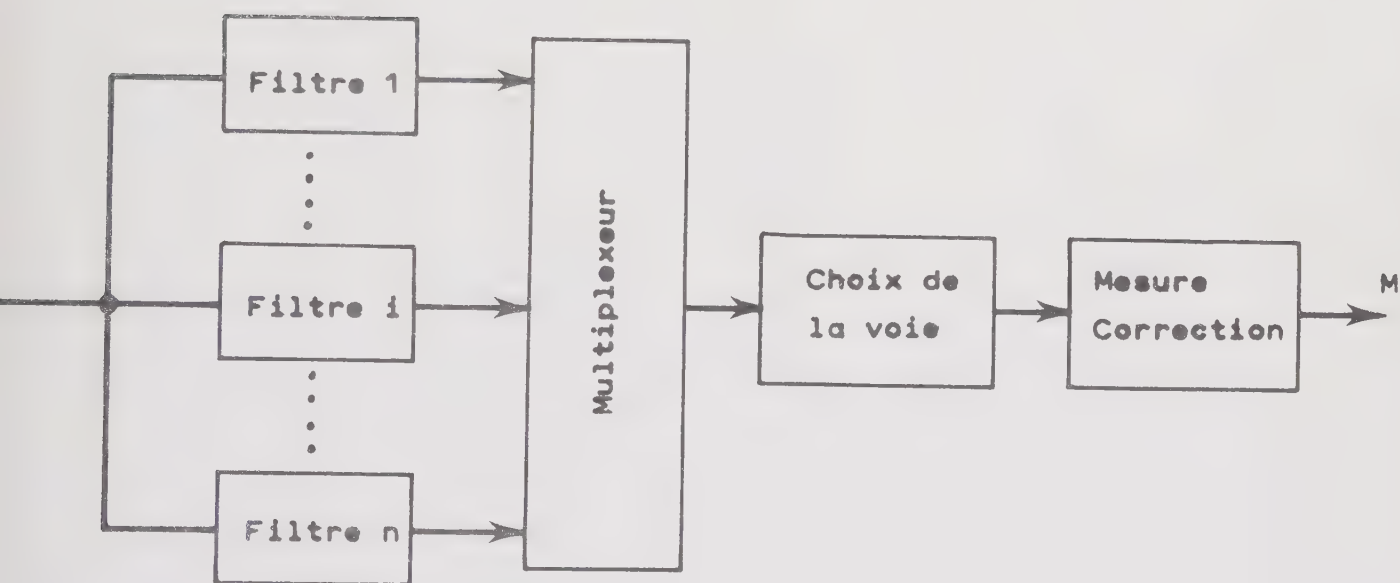


Figure 4 - ANALYSEUR A n FILTRES PASSE-BANDE EN PARALLELE .

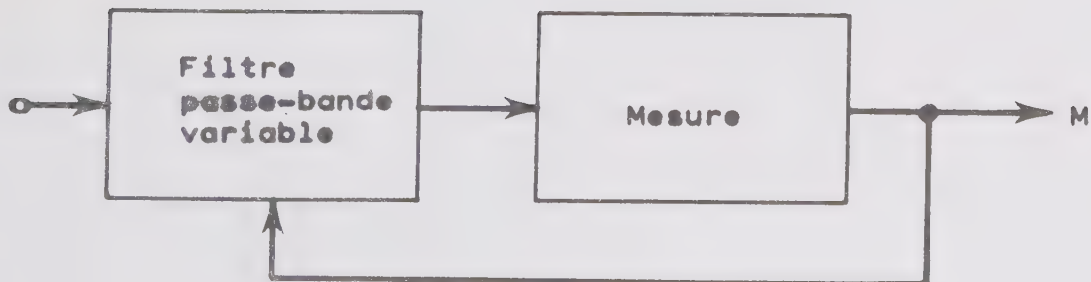


Figure 5 - ANALYSEUR A FILTRE VARIABLE .

La fonction $p(t)$, bien que pouvant être choisie optimum, introduit nécessairement une erreur par rapport au spectre réel de $f(t)$.

Nous distinguons, parmi les analyseurs de mélodie spectraux, des analyseurs :

- à un seul filtre fixe (figure 3)
- à canaux : filtres fixes différents en parallèle avec système de scrutation périodique et choix du canal (figure 4)
- à filtre variable : manuel ou automatique (commande de la fréquence de coupure à l'aide du signal de mesure) (figure 5)
- à calcul direct de $H(\theta, \omega)$ sur ordinateur.

Dans cette dernière catégorie entrent des versions récentes qui utilisent des algorithmes élaborés dérivés de $H(\theta, \omega)$ (cepstre, clipstre, double spectre, product spectrum) qui gagnent en précision et fiabilité, mais ne permettent pas l'analyse en temps réel dans l'état actuel de la technique.

1.4.2.2. *Analyse dans le domaine temporel*

L'analyse de périodicité d'un signal $f(t)$ possédant les mêmes propriétés mathématiques que dans le cas de l'analyse fréquentielle, s'effectue par le calcul de sa fonction d'autocorrélation :

$$\phi(\tau) = \int_{-\infty}^{+\infty} f(t) f(t+\tau) dt$$

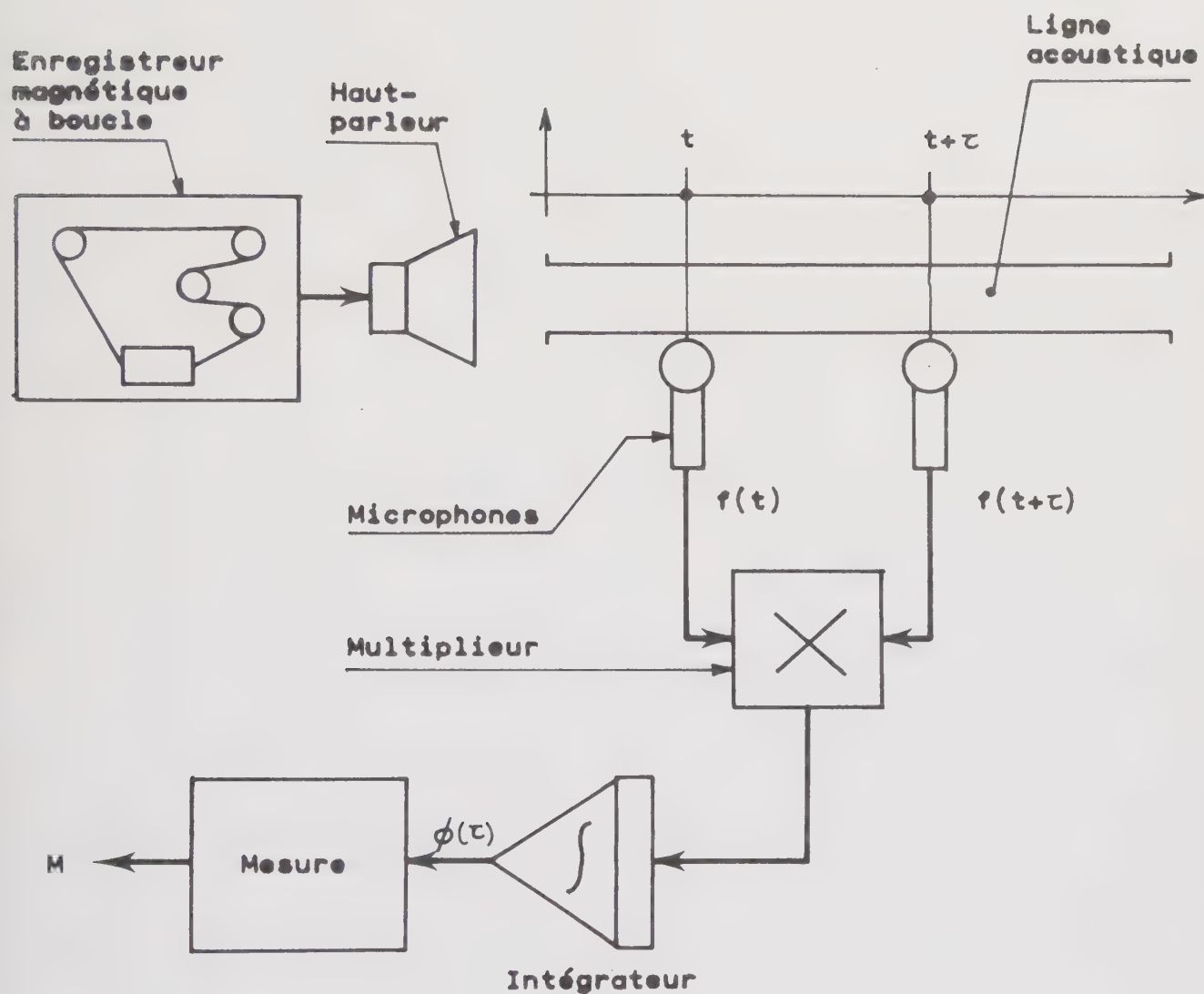


Figure 6 - ANALYSEUR PAR AUTOCORRELATION A LIGNE A RETARD ACOUSTIQUE .

La période d'un signal récurrent est l'intervalle de temps qui sépare l'apparition du premier et du dernier maximum de $\psi(\tau)$.

Pour les mêmes raisons que ci-dessus, on utilise physiquement la fonction d'auto-corrélation quasi-instantanée en pondérant l'intégrale par une fonction $p(t)$ identique à la précédente :

$$\psi(\theta, \tau) = \int_{\theta-T}^{\theta+T} f(t) f(t+\tau) p(t-\theta) dt .$$

La fonction $p(t)$ introduit là aussi une erreur sur l'analyse de $f(t)$.

Les principaux types d'analyseurs à auto-corrélation sont :

- type manuel : dépouillement de l'oscillogramme d'un son de parole
- à ligne à retard (figure 6)
- à calcul direct de $\psi(\theta, \tau)$ ou d'un algorithme dérivé sur un ordinateur.

Il existe une relation étroite entre l'analyse fréquentielle et l'analyse temporelle : la quantité $\Phi(\omega)$ ($\Phi(\omega) = 2\pi |F(\omega)|^2$), appelée densité spectrale d'énergie de $f(t)$ est la partie réelle de la transformée de Fourier de la fonction d'autocorrélation $\psi(\tau)$

1.4.3. Un nouveau type d'analyseur

Le principal obstacle à la réalisation des analyseurs de mélodie suivant les principes décrits ci-dessus, est la difficulté de détecter avec certitude et précision le signal dont le traitement fournira la

valeur de la hauteur cherchée. De plus, pour la plupart, l'utilisation de nombreux filtres introduit des retards, fonctions de la fréquence, difficilement contrôlables. Les calculs de décision et de corrections nécessaires sont souvent confiés à un ordinateur et leur complexité rendent aléatoire, voire même impossible, le fonctionnement en temps réel de l'appareil.

Nous avons pensé qu'il était plus simple et plus logique d'analyser directement le son laryngé qui est le signal source, plutôt que d'essayer de le reconstituer après les déformations qu'il subit dans les cavités du conduit vocal.

Le seul procédé actuel permettant d'obtenir un signal électrique image du mouvement des cordes vocales, dans des conditions d'utilisation simples, est l'électroglottographie introduite par Ph. Fabre en 1957.

Nous avons utilisé cette méthode comme base de réalisation de l'analyseur de mélodie étudié dans les chapitres suivants. Celui-ci possède toutes les caractéristiques fondamentales (précision, fiabilité, fonctionnement en temps réel), alliées à un coût modéré et à une mise en oeuvre aisée.

C H A P I T R E 2

DETECTION ELECTROGLOTTOGRAPHIQUE DU MOUVEMENT DES CORDES VOCALES

L'analyseur de mélodie faisant l'objet de ce mémoire est divisé en deux modules :

- le dispositif de détection du signal image du son laryngé, étudié dans ce chapitre ;
- l'appareil de mesure des intervalles de temps de récurrence des impulsions laryngées décrit au chapitre suivant.

Après un bref aperçu sur les méthodes existantes d'observation du mouvement des cordes vocales, et sur la nature de l'impédance transversale du larynx en haute-fréquence, nous étudions la réalisation d'un électroglottographe suivant le principe introduit par Ph. Fabre.

2.1. LES METHODES D'OBSERVATION DU MOUVEMENT DES CORDES VOCALES

2.1.1. Méthodes d'observation directe

Les premiers essais d'examen des cordes vocales pendant la phonation remontent à Manuel Garcia (1855) qui éclaira et regarda l'intérieur de son propre larynx à l'aide d'un petit miroir de dentiste placé sur la luette et en comprimant fortement sa langue vers le bas pour assurer

un champ visuel d'observation suffisant. Le développement ultérieur des techniques a permis, tout en conservant le système d'observation de Garcia, d'examiner avec plus de précision le mouvement des cordes vocales, en utilisant, soit un éclairage stroboscopique, soit la cinématographie à grande vitesse à l'aide de caméras ultra-rapide.

Ces procédés présentent l'inconvénient majeur de nécessiter une compression vigoureuse de la langue. Il s'ensuit inévitablement une perturbation sur le mouvement des cordes vocales, et l'impossibilité de prononcer clairement d'autres sons que des voyelles tenues. De plus, la mise en oeuvre de telles expériences est très malaisée ; l'exploitation manuelle des résultats est longue et fastidieuse, peu précise dans le cas de la laryngo-stroboscopie qui ne montre qu'un assemblage de petites portions de chaque impulsion laryngée et très coûteuse dans le cas de la cinématographie rapide. Ces énormes limitations d'utilisation ont pratiquement exclu ces méthodes de l'arsenal d'analyse expérimentale courante de la parole.

2.1.2. Méthodes d'observation indirecte

Les travaux de Ph. Fabre (1957) [9][11] ont ouvert la voie à l'observation dynamique aisée des mouvements du larynx et des cordes vocales, par l'introduction de l'électroglottographie.

Le procédé consiste à faire apparaître aux bornes d'électrodes métalliques placées de chaque côté du larynx une tension à haute fréquence modulée en amplitude par les variations d'impédance électrique au niveau de la glotte dues au mouvement des cordes vocales.

Le signal que fournit l'électroglottographe représente avec une bonne approximation la variation d'aire de la glotte. L'appareil présente le gros avantage de ne pas perturber la phonation, car les électrodes, bien placées, n'apportent aucune gêne au locuteur.

Quelques chercheurs [20][21] ont également développé une méthode opto-électronique d'examen du mouvement des cordes vocales. C'est la glottographie photo-électrique. Elle consiste à éclairer fortement le larynx de l'extérieur et par dessous (par transparence) et à recueillir les variations d'éclairement de la glotte à l'aide d'une sonde photo-électrique introduite par la bouche ou par le nez, et placée devant la sortie du larynx. Ce procédé est d'un emploi plus lourd que l'électroglottographie et la sonde peut-être cause de gêne.

Il est cependant à remarquer que le signal issu du glottographe photo-électrique représente directement les variations d'aire de la glotte, dans un plan transversal au larynx, alors que l'électroglottographe mesure les variations d'aire de la surface d'appui des cordes vocales qui se produisent dans un plan vertical. L'utilisation simultanée des deux appareils peut d'ailleurs permettre d'étudier le mouvement des cordes vocales dans les trois dimensions de l'espace et en fonction du temps .

2.1.3. L'électroglottographe de Fabre

Nous avons choisi d'utiliser l'électroglottographe de Fabre comme capteur des impulsions laryngées dans la chaîne d'analyse des variations mélodiques de la voix.

Les raisons de ce choix reposent sur deux considérations :

- l'analyseur de mélodie que nous nous proposons de réaliser doit utiliser un signal image des cycles glottiques, ce qui impose pour le capteur d'être un détecteur du mouvement des cordes vocales.
- des deux appareils existants réalisant cette condition (électroglottographe et glottographe photo-électrique), nous devons utiliser le plus simple à mettre en oeuvre de façon à assouplir l'utilisation du système qui peut devenir ainsi un instrument d'analyse courant d'un laboratoire de phonétique.

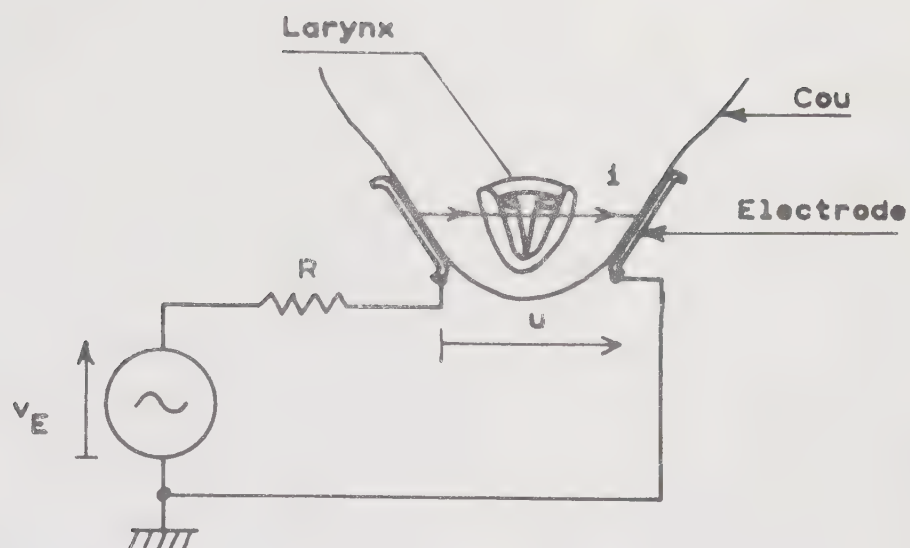


Figure 7 - PRINCIPE DE L'ELECTROGLOTTOGAPHE DE FABRE .

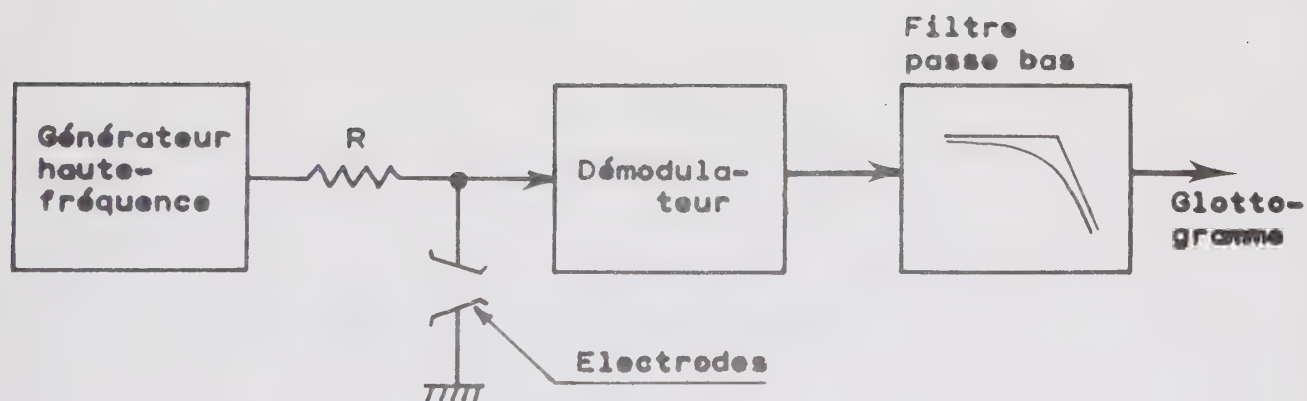


Figure 8 - FONCTIONNEMENT DE L'ELECTROGLOTTOGAPHE .

Le schéma de la figure 7 illustre le principe de fonctionnement de l'électroglottographe de Fabre. Deux électrodes métalliques, enduites de pâte conductrice, placées de part et d'autre du cartilage thyroïde au niveau de la glotte, alimentées par un générateur de tension alternative à travers une résistance (R), permettent le passage d'un courant dans le larynx et présentent entre elles une différence de potentiel dépendant de l'impédance électrique (Z) équivalente à l'ensemble des tissus du cou. Une faible modulation en amplitude apparaît sur la tension inter-électrodes pendant l'émission des sons voisés. Cette modulation est due à l'action de la variation d'impédance électrique du larynx sur les lignes de courant qui traversent la glotte.

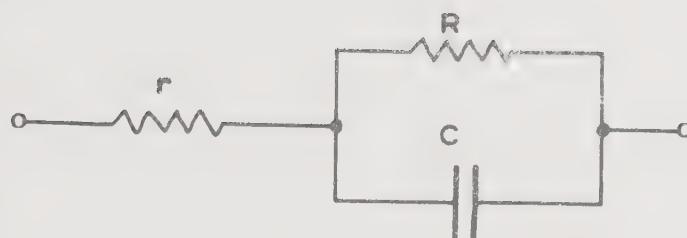
Ph. Fabre a montré que le signal basse fréquence appelé glottogramme, issu de la détection et du filtrage de la tension modulée, est l'image de la variation d'aire de la glotte en fonction du temps.

Le schéma de principe de l'électroglottographe est reproduit à la figure 8

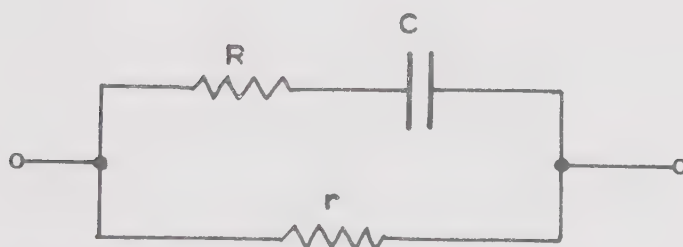
2.2. CARACTERISTIQUES DE L'IMPEDANCE ELECTRIQUE EQUIVALENTE AU LARYNX

L'électroglottographe est un dispositif expérimental d'analyse des mouvements du larynx et des cordes vocales. Le signal électrique fourni renseigne sur la forme de la variation d'aire de la glotte par son amplitude en fonction du temps et sur la hauteur du son laryngé par la durée des cycles glottiques.

Dans le cas de l'analyse de mélodie, c'est essentiellement la seconde caractéristique que nous utilisons. Cependant, il nous a semblé intéressant d'étudier un appareil complet, que la structure modulaire de l'analyseur permettra de faire fonctionner seul.



a) MODELE DE PHILIPSON .



b) MODELE DE FRICKE .

Figure 9 - MODELES DE REPRESENTATION DES
TISSUS BIOLOGIQUES .

La construction d'un électroglottographe selon le principe de Fabre, nécessite la connaissance de la forme et de l'évolution pendant la phonation de l'impédance transversale du larynx en haute fréquence. Nous résumons dans les lignes suivantes les résultats des recherches de M. Duvelloy (1961) 8 qui ont permis d'expliquer le fonctionnement du glottographe et d'interpréter le glottogramme.

2.2.1. Comportement des tissus vivants en présence d'un courant alternatif

La structure des tissus vivants et des assemblages d'organes est compacte, et il est difficile par un examen anatomique de définir des chemins caractéristiques représentables par des impédances connues que pourrait emprunter un courant électrique résultant de l'application en deux points d'une tension alternative. On peut cependant approcher avec une bonne approximation le comportement électrique de ces milieux en utilisant les schémas des modèles de Philippon et de Fricke (figure 5) qui se sont révélés, par l'expérience, être une bonne représentation des tissus dans un domaine de fréquences s'étendant d'une dizaine à quelques centaines de kilohertz.

Les études expérimentales d'impédance des milieux biologiques ont permis de mettre en évidence un phénomène de polarisation interne dû à l'accumulation de charges sur les membranes non conductrices qui oblige à remplacer la capacité C des modèles par une impédance du type $1/\gamma(j\omega)^m$ ($m < 1$) appelée capacité de polarisation.

Il arrive enfin que certains tissus puissent être représentés par une résistance pure, surtout en haute-fréquence, car les valeurs relatives des éléments du schéma peuvent permettre de négliger la capacité.

2.2.2. Schéma équivalent au larynx au repos

Par l'examen anatomique de la coupe du cou au niveau de la glotte, et la mesure de l'impédance complexe du larynx

à différentes fréquences, M. Duvelleroy a pu mettre au point un modèle représentant avec une bonne approximation le comportement en courant alternatif des tissus compris entre les deux électrodes (figure 10).

Les expériences ont été réalisées à l'aide d'électrodes en argent de $0,8 \text{ cm}^2$ appliquées de part et d'autre du cartilage thyroïde après destruction de l'épiderme par érosion à la pâte conductrice et interposition de papiers filtre imbibés d'une solution de NaCl, de façon à minimiser la résistance série due aux plans cutanés.

Le dispositif de mesure d'impédance est schématisé à la figure 12. On applique à l'entrée de l'amplificateur vertical d'un oscilloscope la tension prélevée aux bornes des électrodes et à l'entrée de l'amplificateur horizontal cette même tension déphasée de $\pi/2$; en synchronisant l'allumage du spot sur les passages par zéro de la tension d'excitation V_E , le point apparaissant sur l'écran représente l'impédance inconnue (Z) dans le plan complexe initialement étalonné.

Duvelleroy a pu déduire de ses expériences les ordres de grandeur relatifs des impédances composant le modèle du larynx, ce qui a permis d'en simplifier le schéma (figure 11). Il apparaît en particulier que la capacité glottique (C) varie de façon considérable (0 à 50 nF valeur typique) pendant la phonation, mais elle est shuntée par une résistance très faible qui en annule l'effet.

2.2.3. Etude des variations de l'impédance du larynx au cours de la phonation

Les variations d'impédance du larynx dues au mouvement des cordes vocales sont très faibles : la variation relative du module est de l'ordre de 0,1 %. Leur étude expérimentale n'a pu être réalisée qu'à l'aide d'une méthode différentielle utilisant un pont d'impédances

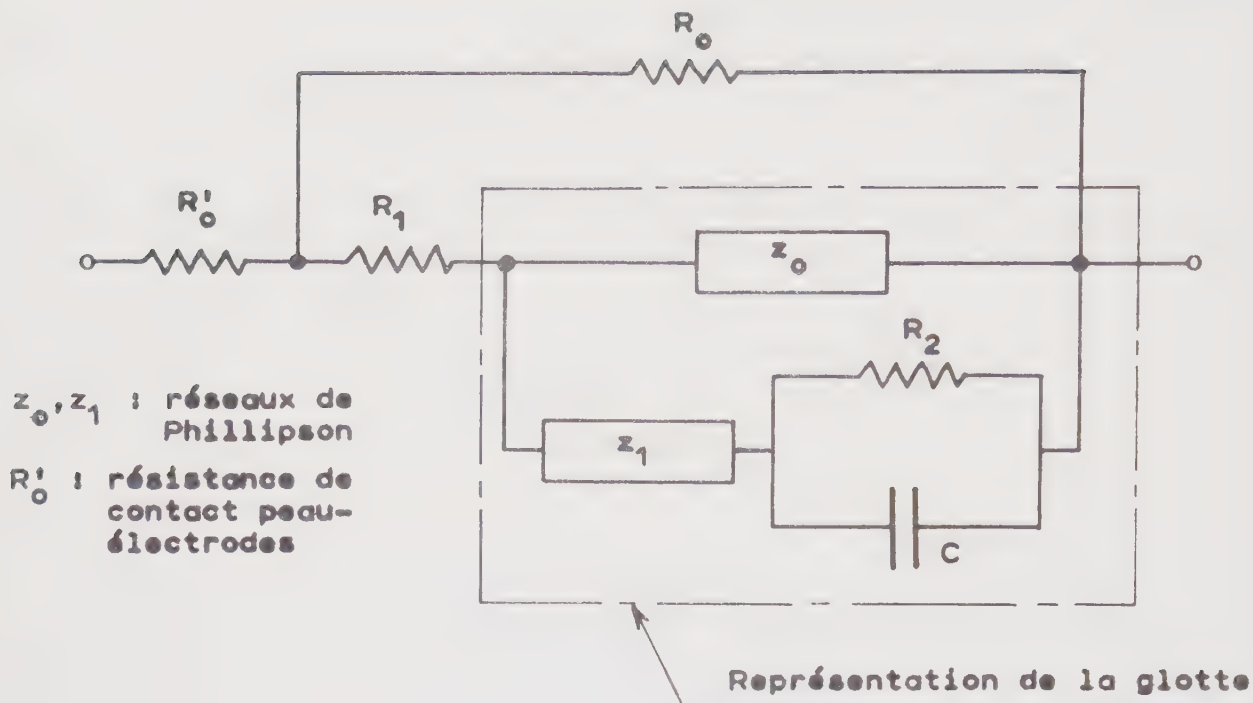


Figure 10 - SCHEMA EQUIVALENT AU LARYNX AU REPOS .
(D'APRES DUVELLEROY)

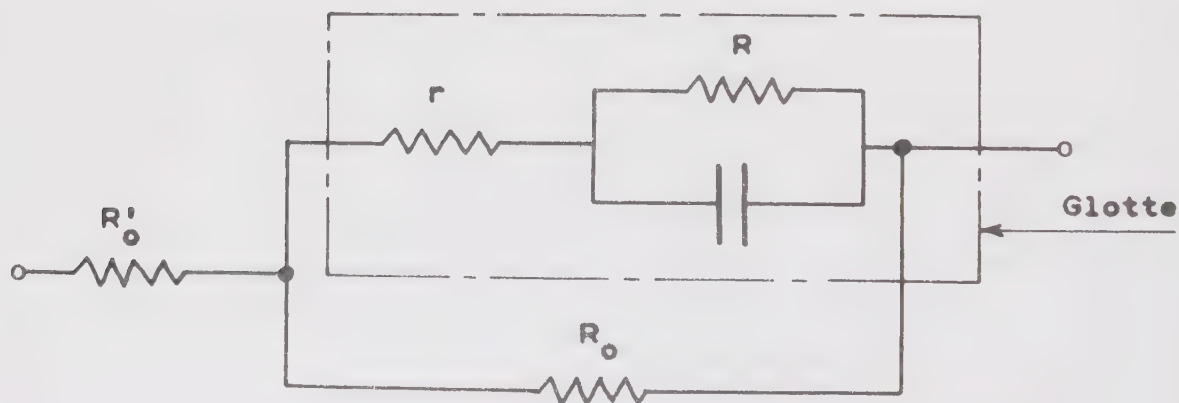


Figure 11 - SCHEMA EQUIVALENT AU LARYNX SIMPLIFIE .
(D'APRES DUVELLEROY)

(figure 16) dont deux branches sont constituées par deux résistances pures et les autres branches par l'impédance (Z) à analyser d'une part, et un réseau d'équilibrage à résistance et capacité réglables en parallèle d'autre part.

La tension ΔV apparaissant aux bornes de la diagonale AB du pont, est proportionnelle, si l'impédance de charge est très grande, à la variation relative d'impédance qui vaut :

$$\frac{\Delta Z}{Z} = \frac{\Delta Z_0}{Z_0} + j \Delta \phi \quad (Z = Z_0 e^{j\phi})$$

La partie réelle de ΔV représente donc la variation relative du module de Z et la partie imaginaire la variation absolue de phase.

Les variations de l'impédance mesurée entre les électrodes au cours de la phonation ont deux causes essentielles :

- les déformations relativement lentes et de grande amplitude des tissus entourant le larynx, dues aux mouvements d'ensemble inévitables de celui-ci (fréquences inférieures à 10 Hz) ;
- la modification de la configuration interne du larynx, beaucoup plus rapide, due à l'écartement et à l'accroissement des cordes vocales (80 à 600 Hz).

Les différences de vitesse des deux types de variations d'impédance permettent de les dissocier et de ne s'intéresser qu'à celle qui représente le mouvement des cordes vocales.

Les éléments susceptibles de varier, dans le modèle électrique de la glotte (figure 17) sont la capacité glottique (dont les cordes vocales forment schématiquement les armatures et l'air le diélectrique) et la résistance glottique série (incluse dans z_1).

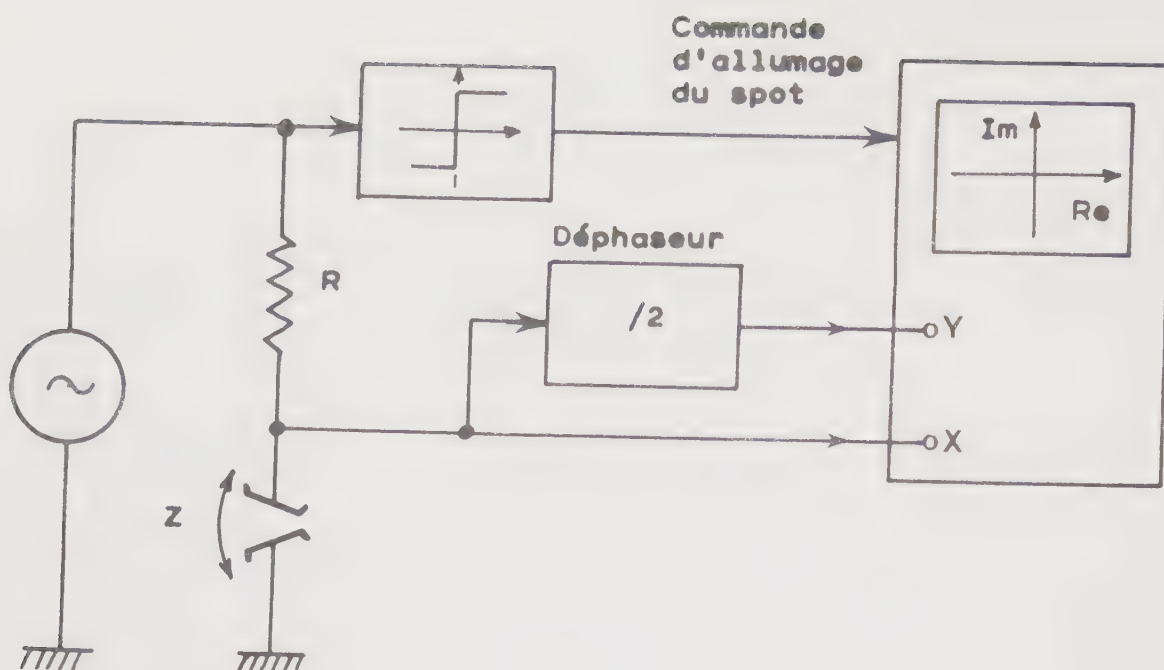


Figure 12 - DISPOSITIF DE MESURE D'IMPEDANCE EN H.F.
(D'APRES DUVELLEROY)

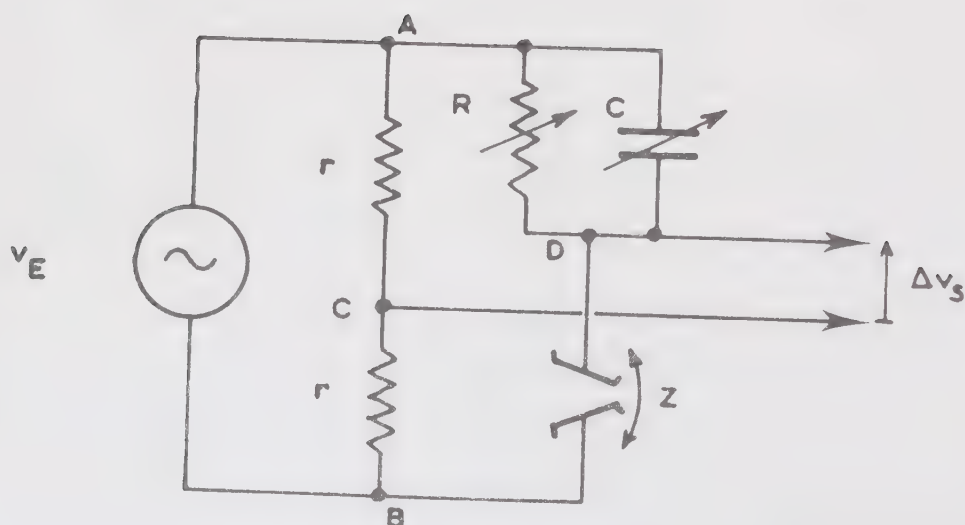


Figure 13 - MESURE DE VARIATIONS D'IMPEDANCE
A L'AIDE D'UN PONT .

Nous avons vu que la capacité glottique, shuntée par une résistance très faible, n'intervient pas dans l'impédance du larynx. Duvelleyroy a pu confirmer ce fait en montrant la concordance entre les valeurs théoriques et expérimentales des variations de l'impédance du larynx dues à la phonation.

2.2.4. Conclusion

Les mouvements des cordes vocales ne provoquent pratiquement qu'une variation de la résistance série du schéma équivalent de la glotte, en voix de poitrine (registre normal), ce qui permet d'interpréter l'amplitude du glottogramme comme étant l'image de l'écartement instantané des cordes vocales.

Duvelleyroy a également montré que l'ouverture de la glotte correspond à une diminution de la résistance glottique.

2.3. INTERPRETATION DU GLOTTOGRAMME

Le glottogramme en régime permanent est un signal de forme générale et d'amplitude quasi stationnaires rappelant la forme d'une oscillation de relaxation.

D'après les conclusions de Duvelleyroy, la modulation par le mouvement des cordes vocales de la tension aux bornes des électrodes a la forme représentée à la figure 14 : on y remarque l'effet de diminution de résistance glottique lors de l'écartement des cordes vocales (représenté par une variation brusque).

Le glottogramme, issu de la détection de cette modulation, est cependant généralement présenté conformément à la figure 15 qui est ainsi le diagramme en fonction du temps de la distance instantanée séparant les cordes vocales. La partie à pente positive (O) est l'ouverture de la glotte, la partie à pente négative (F) la fermeture et le palier (A) la phase d'accollement dont la pente est rarement nulle, car elle représente l'aire

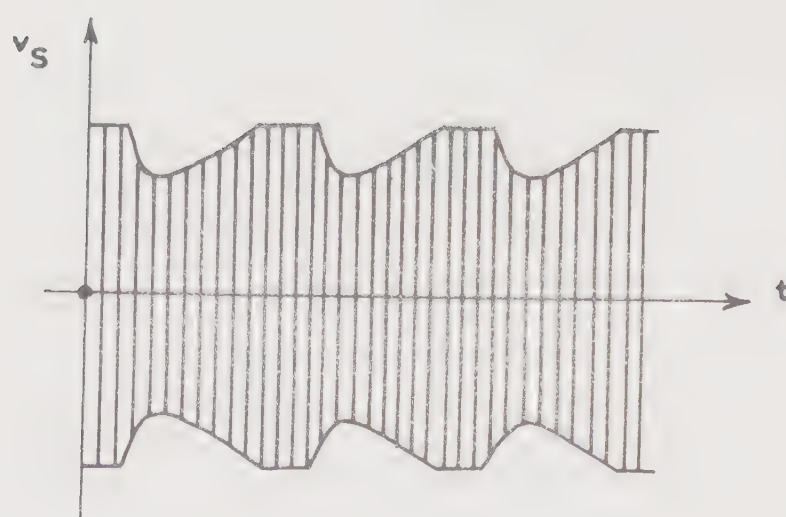


Figure 14 - TENSION H.F. MODULEE AUX BORNES DES ELECTRODES .

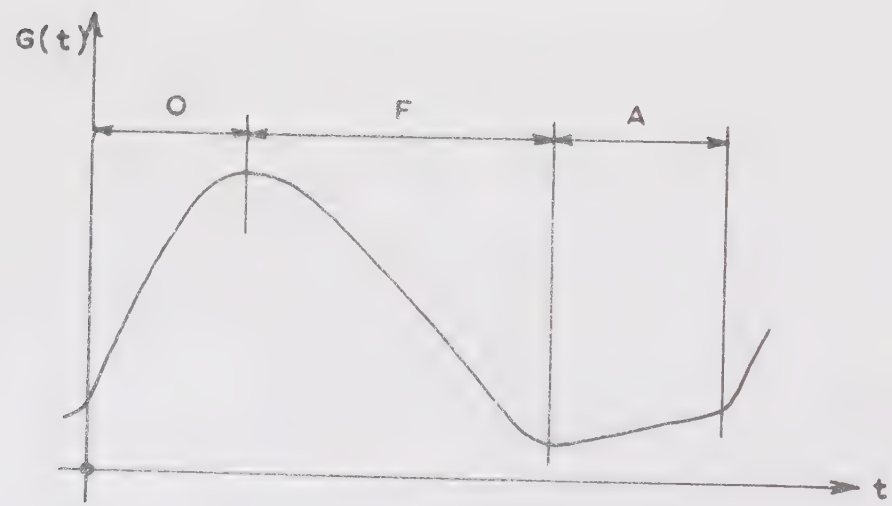


Figure 15 - ALLURE D'UN GLOTTOGRAMME .

instantanée des surfaces des cordes vocales en contact qui peuvent être plus ou moins appuyées.

L'amplitude du glottogramme et les valeurs relatives des durées des trois phases du cycle glottique varient suivant la configuration du conduit vocal et la conduite phonatoire adoptée. Elle est surtout sensible aux variations de hauteur, d'intensité, d'impédance acoustique ramenée sur le larynx par les cavités sus-glottiques et aux mouvements verticaux du larynx pendant la phonation. Ph. Fabre et R. Husson ont analysé et répertorié les modifications correspondantes du glottogramme [10] [25]

Le temps T nécessaire à l'exécution des trois phases du cycle glottique est l'intervalle de temps séparant deux impulsions laryngées ; il est à l'origine de la sensation de hauteur à l'audition de la parole. La mesure de T en fonction du temps fournira donc la "courbe mélodique" du signal de parole émis.

Il est également intéressant de remarquer que le spectre fréquentiel du glottogramme montre une décroissance de l'amplitude des harmoniques de 12 dB par octave.

2.4. ETUDE ET REALISATION D'UN ELECTROGLOTTOGRAPHE

L'électroglottographe est un appareil qui effectue la mesure des variations rapides du module de l'impédance instantanée du larynx. Il est composé de trois circuits élémentaires ayant chacun une fonction propre, placés en cascade :

- le dispositif de modulation de la tension haute-fréquence fournie par le générateur interne ; il est constitué d'un circuit passif associé à l'impédance à analyser
- le circuit de démodulation d'amplitude détectant le signal image du mouvement des cordes vocales
- le banc de filtres permettant d'affiner ce signal.

2.4.1. Dispositif de modulation en amplitude de la tension porteuse

2.4.1.1. Principes et montages possibles

Les variations d'une impédance dues à des modifications de sa structure interne, peuvent être mises en évidence quantitativement par comparaison à un réseau d'impédances fixes et connues. Les montages utilisés de façon classique dans les circuits de mesure d'impédance sont notamment :

- le montage en pont
- le montage potentiométrique.

Le montage en pont (figure 16-a) permet de réaliser une mesure différentielle faisant apparaître directement les variations de l'impédance inconnue.

Un générateur de tension alternative haute-fréquence alimente un réseau comportant deux branches ; la première est composée de deux résistances pures égales et la seconde est matérialisée par une impédance complexe réglable (circuit résistance capacité parallèle) en série avec l'impédance à analyser. La sortie du montage est constituée par la tension V_s prélevée entre les points milieux des deux branches. L'équilibre étant préalablement réalisé en égalant (Z') à (Z) en module et argument, on montre qu'une petite variation ΔZ de (Z) produit une tension V_s proportionnelle à la variation relative d'impédance :

$$v_S = \frac{V_E}{4} \frac{\Delta Z}{Z} \Rightarrow v_S = \frac{V_E}{4} \left| \frac{\Delta Z}{Z} \right|$$

$$\text{où } \frac{\Delta Z}{Z} = \frac{\Delta Z_0}{Z_0} + j \Delta \phi \quad (Z = Z_0 e^{j\phi})$$

Notations : Z impédance à mesurer

$|Z|$ module de Z

ΔZ variation glottique de Z

ϕ argument de Z

V_S, V_E amplitudes des tensions v_S et v_E .

La variation ΔZ étant très petite devant (Z) , on admet que V_S est proportionnelle à $|\Delta Z|$.

La sensibilité du montage (rapport de la tension de sortie à la variation d'impédance) ne dépend que de l'amplitude V_E de la tension délivrée par le générateur.

Ce montage présente un grave inconvénient : l'instabilité de la valeur moyenne de l'impédance du larynx due à des mouvements lents de celui-ci nécessite un ajustage fréquent du pont; cet ajustage est long et délicat en raison des deux paramètres à modifier simultanément ; de plus, le point de fonctionnement doit être légèrement décalé par rapport à l'équilibre pour éviter une distorsion dite de "doublement de fréquence" (figure 17), ce qui a pour effet de diminuer la sensibilité du circuit.

Le montage potentiométrique (figure 18 b) est constitué d'une résistance pure R en série avec l'impédance inconnue (Z) . La tension haute-fréquence du générateur alimente l'ensemble. L'amplitude et la phase de la tension v_S aux bornes de (Z) comparées à celles de la tension v_E du générateur permettent de déterminer le module et l'argument de l'impédance analysée :

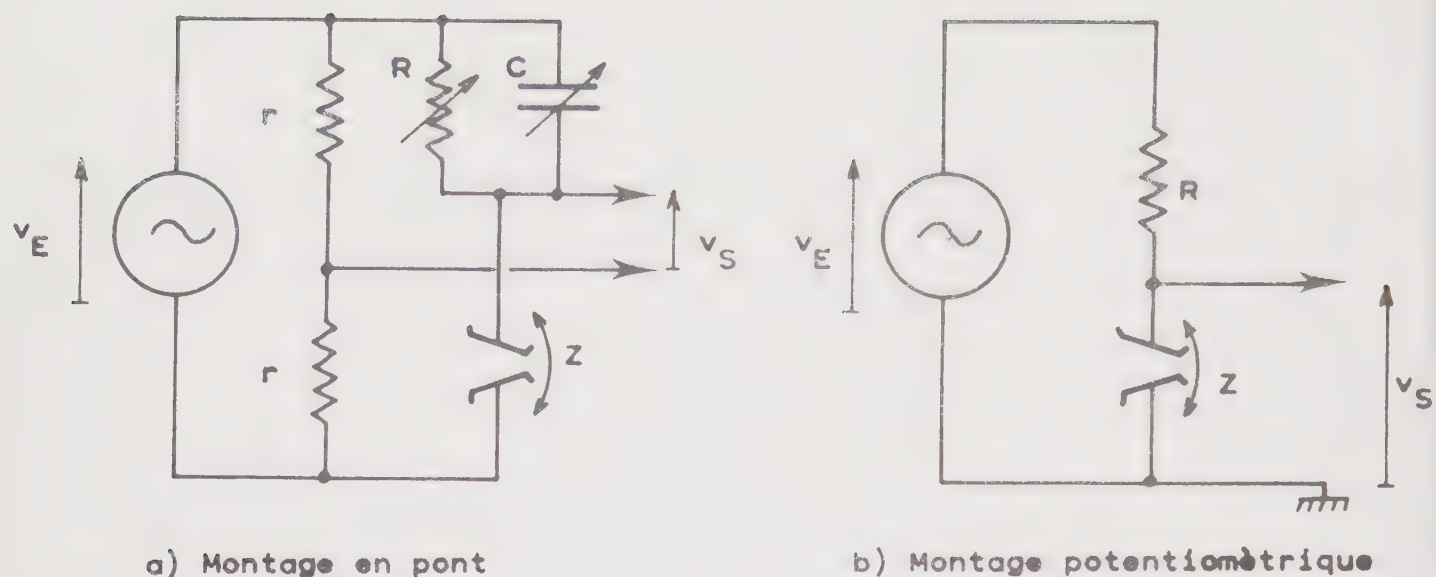


Figure 16 - DIFFERENTS MONTAGES DE MESURE D'IMPEDANCE .

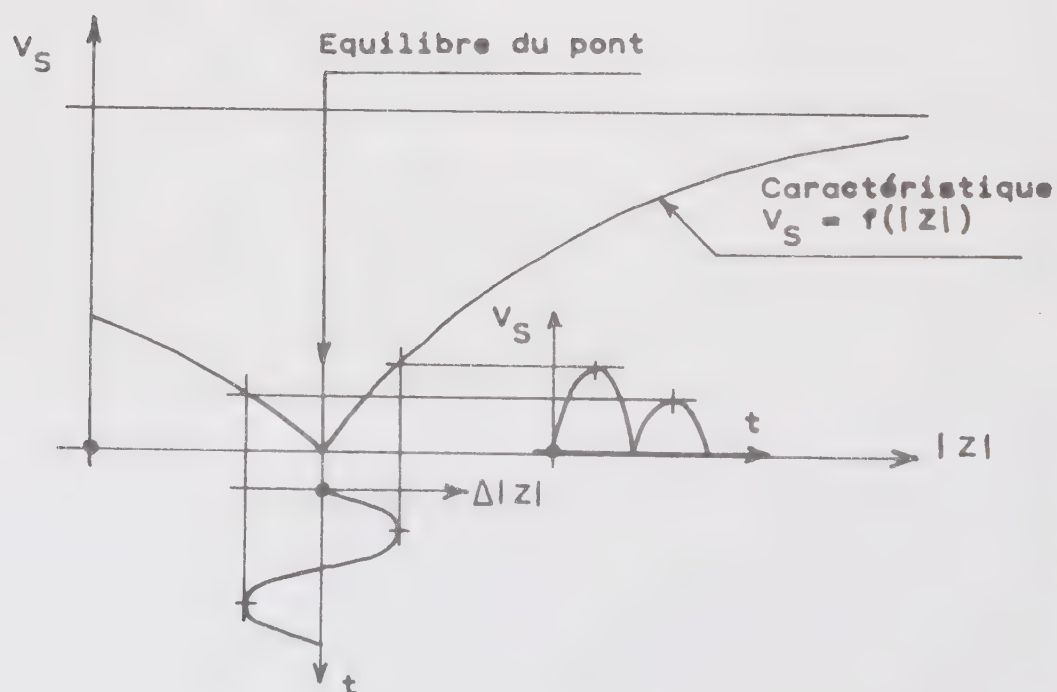


Figure 17 - "DOUBLEMENT DE FREQUENCE" AU POINT D'EQUILIBRE DU PONT .

le montage potentiométrique réalise une mesure directe.

Les variations d'amplitude ΔV_S de V_S peuvent être déterminées en fonction des variations du module de (Z) :

$$V_S = \frac{Z}{R+Z} V_E \Rightarrow V_S = \frac{|Z|}{|Z+R|} V_E$$

ϕ désignant l'argument de (Z) , on a :

$$V_S = \frac{|Z|}{(R^2 + |Z|^2 + 2R|Z|\cos\phi)^{1/2}} V_E$$

de cette expression, on déduit :

$$\frac{\Delta V_S}{V_S} = \frac{1 + \frac{|Z|}{R} \cos\phi}{1 - \left(\frac{|Z|}{R}\right)^2 + 2\frac{|Z|}{R} \cos\phi} \frac{\Delta|Z|}{|Z|} = k \cdot \frac{\Delta|Z|}{|Z|}$$

Comme pour le montage en pont la faible valeur de $\Delta|Z|$ par rapport à $|Z|$ permet de dire que ΔV_S est proportionnelle à la variation d'impédance.

La quantité k représentant la sensibilité du montage, est décroissante en $(|Z|/R)$ et inférieure à 1. Il est donc intéressant de choisir la résistance R grande devant $|Z|$ pour augmenter l'indice de modulation de la tension V_S . Mais il s'en suit une diminution de l'amplitude de V_S qu'il faut compenser par une amplification; le choix de R et du gain de l'amplificateur résulte donc d'un compromis.

2.4.1.2. Etude du montage adopté

Les essais que nous avons effectués avec les deux circuits nous ont confirmé que le montage

en pont est d'un emploi délicat en raison de son instabilité et de ses réglages longs et difficiles. De plus, sa sensibilité n'est pas meilleure que celle du montage potentiométrique.

C'est donc ce dernier que nous avons choisi pour produire la tension haute-fréquence modulée en amplitude par les variations d'impédance du larynx.

Afin de choisir la valeur de la résistance R placée en série avec le larynx, nous avons déterminé l'ordre de grandeur du module de Z dans les conditions normales de la glottographie. Le principe de la mesure est de calculer $|Z|$ à partir de l'amplitude V_S et de la phase ψ de v_S par rapport à la tension v_E du générateur dans le montage potentiométrique (la résistance R étant calibrée) :

$$|Z| = \frac{R V_S}{\sqrt{V_E^2 + V_S^2 - 2 V_E V_S \cos \psi}}.$$

Le choix de la fréquence de la porteuse n'est pas critique. La résistance série équivalente à la peau diminuant avec la fréquence, celle-ci doit être prise suffisamment grande, ce qui permet également une bonne définition et un filtrage ultérieurs aisés de la modulation. Son influence sur la sensibilité du glottographe est négligeable. Les valeurs couramment employées dans les réalisations antérieures varient de 200 à 500 KHz ; l'utilisation de fréquences plus élevées n'entraîne que des complications de câblage inhérentes à la manipulation des hautes-fréquences. Le glottographe que nous avons réalisé utilise une porteuse de 500 KHz.

Le contact au niveau des électrodes doit être le plus étroit possible. L'utilisation d'une pâte conductrice réduit la résistance de contact, la valeur de $\frac{\Delta Z}{Z}$ étant ainsi augmentée.

Les valeurs du module de l'impédance du larynx que nous avons mesurées varient entre 40 Ω et 200 Ω . Il est évident que plus cette valeur est élevée (en particulier à cause de la présence de tissus adipeux), plus la sensibilité diminue, car la variation relative d'impédance est réduite. Nous pourrions donc choisir pour R une résistance variable de 2000 Ω qu'il sera nécessaire d'ajuster pour chaque locuteur au $1/10$ de la valeur de $|Z|$ de façon à garantir une tension de sortie V_S ayant toujours le même ordre de grandeur pour respecter les conditions d'amplitude d'entrée de l'étage suivant. Le choix de $R = 10 |Z|$ permet d'obtenir, à 500 KHz, une sensibilité $k = 0,9$, qui est acceptable.

Le générateur de tension haute-fréquence que nous avons utilisé dans l'électroglottographe est un simple oscillateur à un transistor, de type Hartley, suivi d'un étage collecteur commun abaisseur d'impédance. Il est accordé sur 500 KHz et délivre une tension de 8 Volts crête-à-crête. Sa stabilité en fréquence et en amplitude est suffisante et ne nécessite aucune correction.

La figure 12 montre le schéma de réalisation de l'ensemble générateur-modulateur.

Nous avons modifié le montage⁴⁵ potentiométrique classique en insérant un transformateur⁴⁶ qui joue le rôle d'adaptateur d'impédances entre R et Z. Cette disposition a été également utilisée dans la réalisation de Frokjaer-Jensen .

Le transformateur est constitué d'un primaire (n_1 spires), d'un secondaire (n_2 spires) et d'un petit enroulement (n_3 spires) fermé sur l'impédance Z . En considérant que le secondaire est chargé par une impédance infinie, le générateur "voit" aux bornes du primaire une impédance $Z_p = Z/k_1^2$ (où $k_1 = n_3/n_1$) (figure 19).

La tension de sortie v_s du montage, disponible au secondaire vaut alors :

$$v_s = k_2 v_p = \frac{Z}{Z + R' k_1^2} k_2 E \quad (k_2 = n_2/n_1)$$

Ce dispositif est identique au montage potentiométrique où $R = R' k_1^2$. Ceci permet, en choisissant $k_1 < 1$, de diminuer la puissance demandée à l'oscillateur en le chargeant par une résistance R' de forte valeur sans altérer sa sensibilité du système aux variations de Z . De plus, en utilisant un rapport de transformation k_2 supérieur à l'unité, il est possible de produire une préamplification qui permettra de réduire le gain de l'amplificateur en aval, autorisant ainsi l'utilisation d'un simple transistor.

2.4.2. Détection de la modulation d'amplitude

2.4.2.1. *But et principe de la détection*

La démodulation d'amplitude a pour but de mettre en évidence une grandeur électrique (courant ou tension) représentant les variations, relativement lentes, de l'amplitude d'un signal à fréquence élevée.

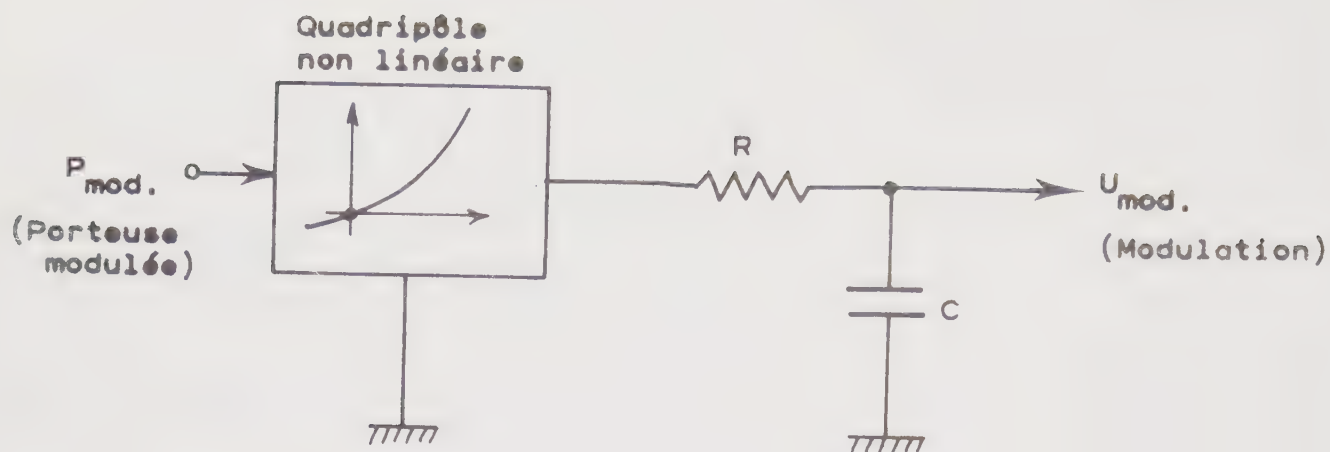


Figure 20 - PRINCIPE DU DETECTEUR DE MODULATION D'AMPLITUDE .

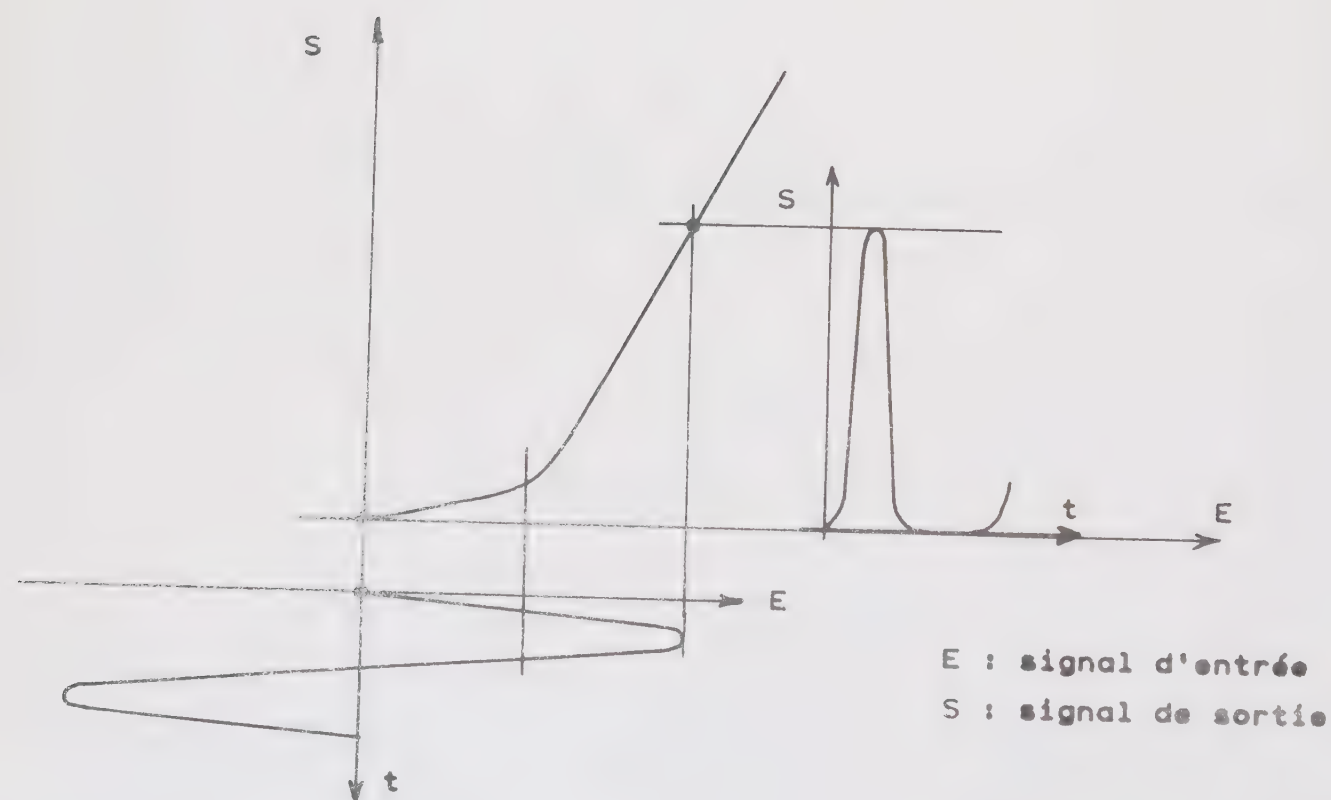
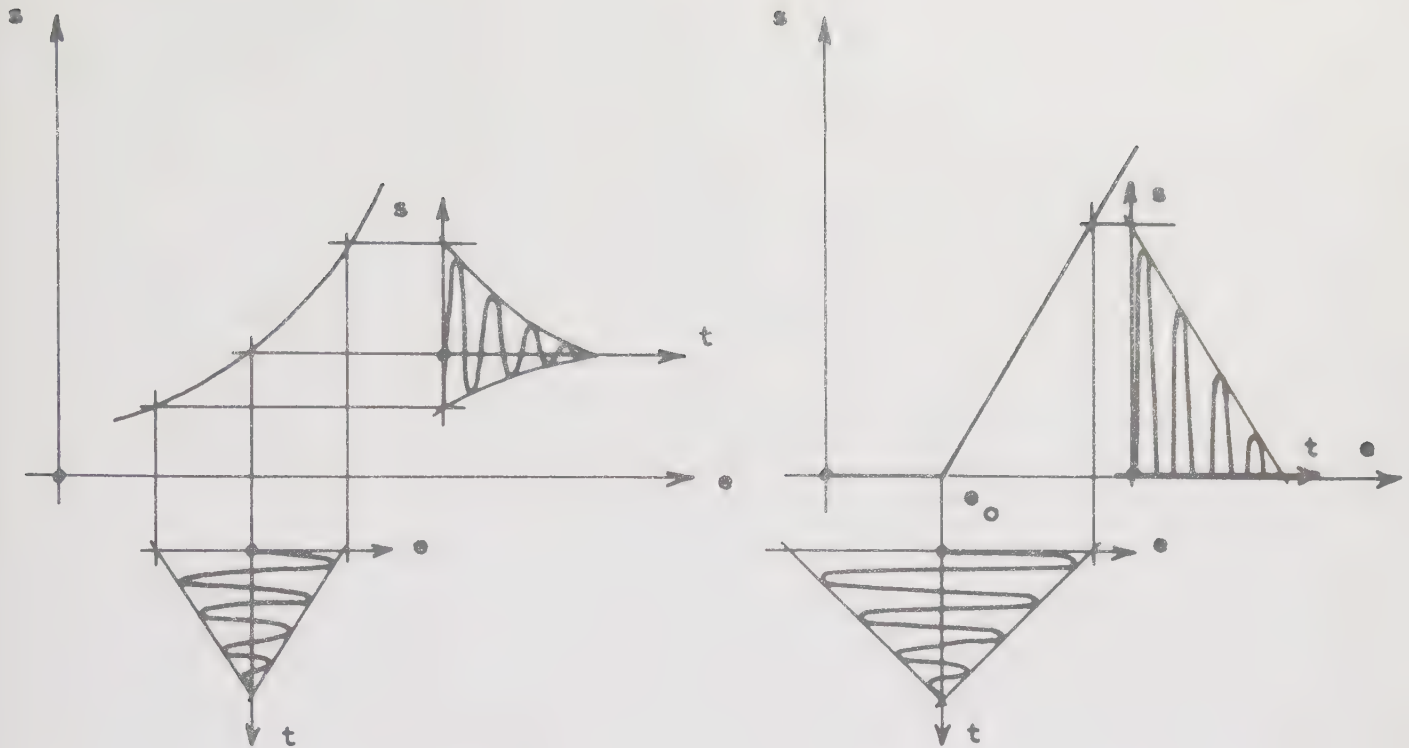


Figure 21 - CARACTERISTIQUE D'UN DETECTEUR A SEMI-CONDUCTEUR .



a) Détection quadratique
(faible niveau)

b) Détection linéaire
(fort niveau)

Figure 22 - TYPES DE DETECTION SUIVANT LE NIVEAU DE LA PORTEUSE .

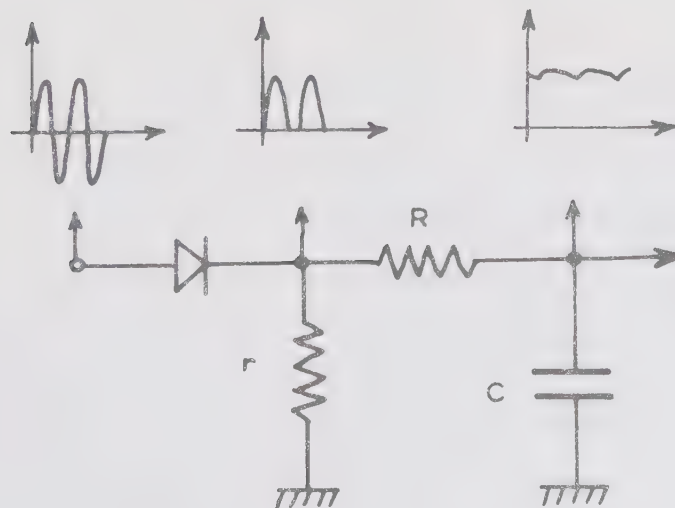


Figure 23 - DETECTEUR SIMPLE A DIODE .

Le principe d'un détecteur de modulation d'amplitude (figure 20) consiste à rendre inégales les alternances positives et négatives de l'onde porteuse à l'aide d'un quadripôle à caractéristique de transfert non linéaire, de façon à faire apparaître une valeur moyenne de signal non nul variant comme l'amplitude de l'onde incidente. La valeur moyenne du signal ainsi détecté est finalement extraite à l'aide d'un circuit de filtrage passe-bas.

Les éléments non linéaires actuellement utilisés dans les détecteurs fonctionnant aux fréquences qui nous intéressent (inférieures à 1 MHz) sont les diodes et les transistors à semi-conducteurs. Les caractéristiques de transfert de ces dispositifs sont semblables à la figure 21.

La linéarité de la détection réalisée à l'aide de ces composants dépend de l'amplitude du signal haute-fréquence à traiter et de l'importance de la modulation : une porteuse de faible amplitude (de l'ordre de la tension de seuil) sera détectée suivant une loi quadratique et une porteuse à fort niveau (supérieur à 1 V) suivant une loi linéaire (figure 22). C'est ce dernier cas qui correspond au signal issu du dispositif de modulation de l'électroglottographe.

2.4.2.2. *Détection des signaux forts à l'aide d'une diode*

Le schéma de principe du détecteur à diode le plus simple (figure 23) permet d'en comprendre aisément le fonctionnement en présence de signaux forts.

Les alternances positives du signal haute-fréquence incident polarisent la diode dans le sens passant au-delà du seuil de conduction et apparaissent aux bornes de la résistance (r), alors que les alternances négatives la polarisent dans le sens bloquant, ce qui annule le courant dans (r).

L'amplitude du signal basse-fréquence obtenu à la sortie de la cellule de filtrage (R-C) est proportionnelle à la modulation de la porteuse, la linéarité étant satisfaisante pour les portions de caractéristique éloignées du seuil de conduction.

Les détecteurs à diode, dont le montage est très simple, sont utilisés lorsque l'onde porteuse, d'amplitude suffisamment grande, est fortement modulée. Si l'indice de modulation est très faible, comme c'est le cas en électroglottographie, on a recours aux systèmes à transistors, dont la fonction de détection est identique à celles des diodes, mais qui présentent des avantages supplémentaires.

2.4.2.3. *Circuits détecteurs à transistors*

Les détecteurs de modulation d'amplitude à transistors permettent d'extraire de très faibles modulations de la porteuse grâce à leur propriété d'amplification. Les types de transistors utilisés dans ces circuits, à jonction ou à effet de champ, conduisent tous au même schéma de principe (figure 24).

Le fonctionnement du détecteur à transistor est illustré par le schéma fonctionnel de la figure 25. Il se décompose en quatre phases :

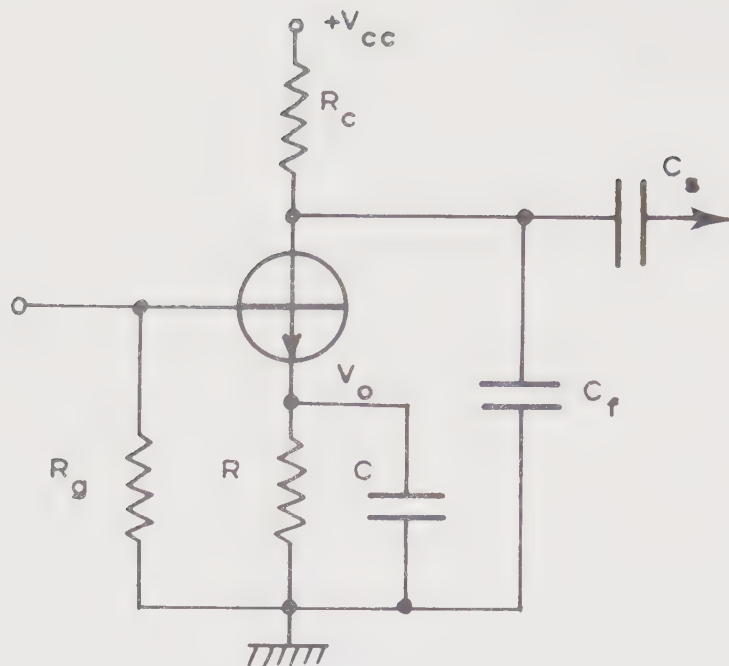


Figure 24 - DETECTEUR A TRANSISTOR .

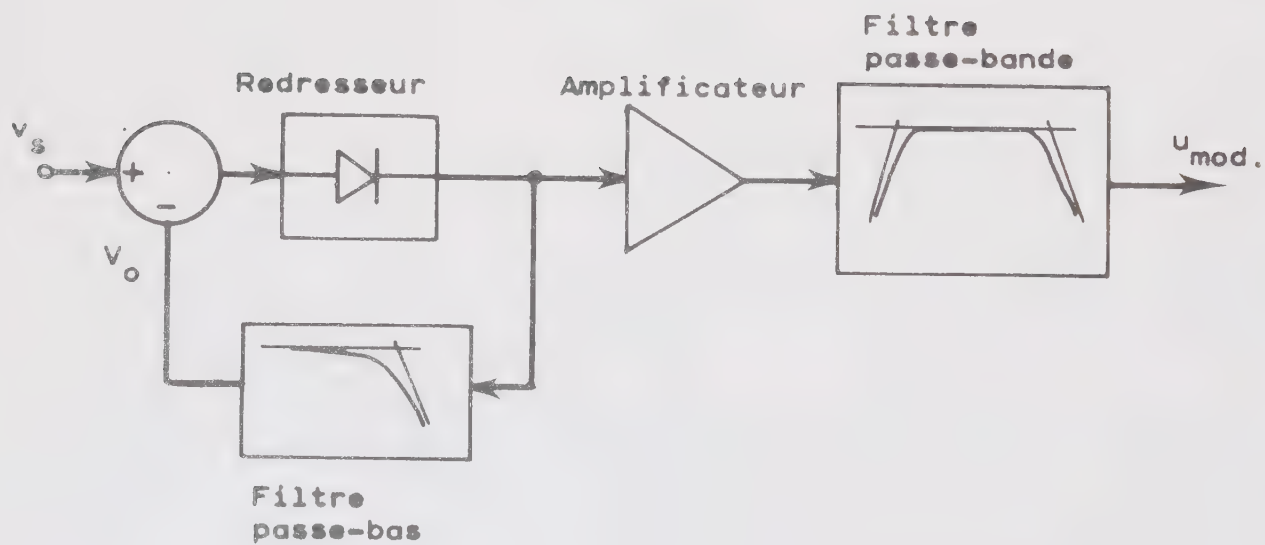


Figure 25 - PRINCIPE DE FONCTIONNEMENT DU DETECTEUR A TRANSISTOR .

- redressement de l'onde porteuse à l'aide de la diode base-émetteur ;
- soustraction au signal d'entrée d'une tension continue V_O fonction de l'amplitude du signal d'entrée et filtrée par le réseau R-C du circuit d'émetteur dont l'impédance, faible pour les valeurs supérieures à $f_C = 1/2\pi RC$ est grande pour les valeurs inférieures ;
- amplification des composantes de $(V_E - V_O)$ de fréquence supérieure à f_C , grâce à la contre-réaction sélective réalisée par le filtre (R-C) ; le dispositif se comporte donc comme un amplificateur des crêtes de la porteuse (figure 26) ;
- filtrage passe bas ; il est constitué par la capacité C_f qui élimine les résidus haute-fréquence du signal de sortie.

Le détecteur que nous avons inclus dans l'électroglottographe est réalisé avec un transistor à jonction au silicium.

La limitation d'amplitude à 2,5 V en entrée nous a fait choisir, pour la tension v_S issue du dispositif de modulation, une amplitude de 1,5 V autour de laquelle le gain du transistor est acceptable. Ceci nécessite d'introduire un amplificateur entre la sortie du circuit modulateur et l'entrée du détecteur. L'impédance d'entrée de cet amplificateur doit être très grande pour ne pas charger le secondaire du transformateur et sa bande passante doit s'étendre au-delà de 500 KHz ; ces conditions peuvent être réalisées à l'aide d'un transistor à effet de champ (figure 27).

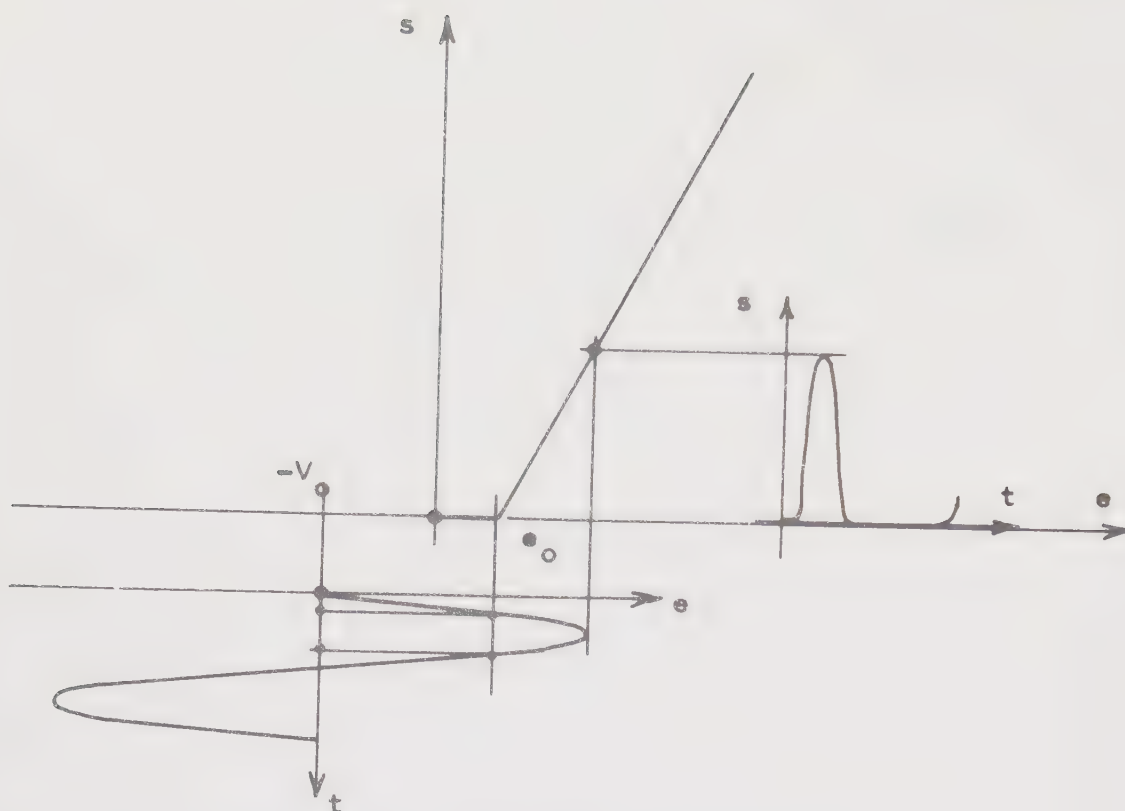


Figure 26 - AMPLIFICATION DES SOMMETS POSITIFS DE LA PORTEUSE .

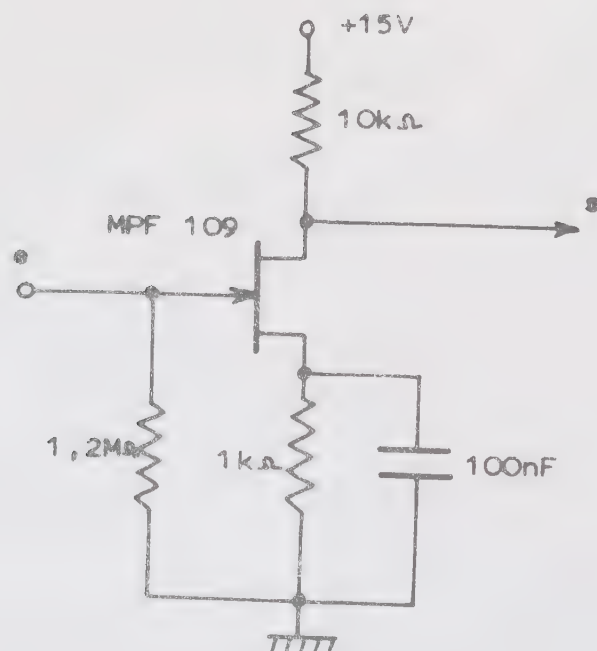


Figure 27 - AMPLIFICATEUR A TRANSISTOR A EFFET DE CHAMP .

A la sortie du glottographe, on peut désirer disposer de deux signaux : l'un représente uniquement les mouvements des cordes vocales (voie "alternative") alors que l'autre fournit également les variations lentes d'impédance dues aux mouvements d'ensemble du larynx pendant la phonation ("voie continue").

Le détecteur de la voie alternative (figure 28 a) est calculé pour atténuer les composantes de modulation de fréquence inférieure à 10 Hz (réseau R-C de constante de temps $\tau = 100$ ms).

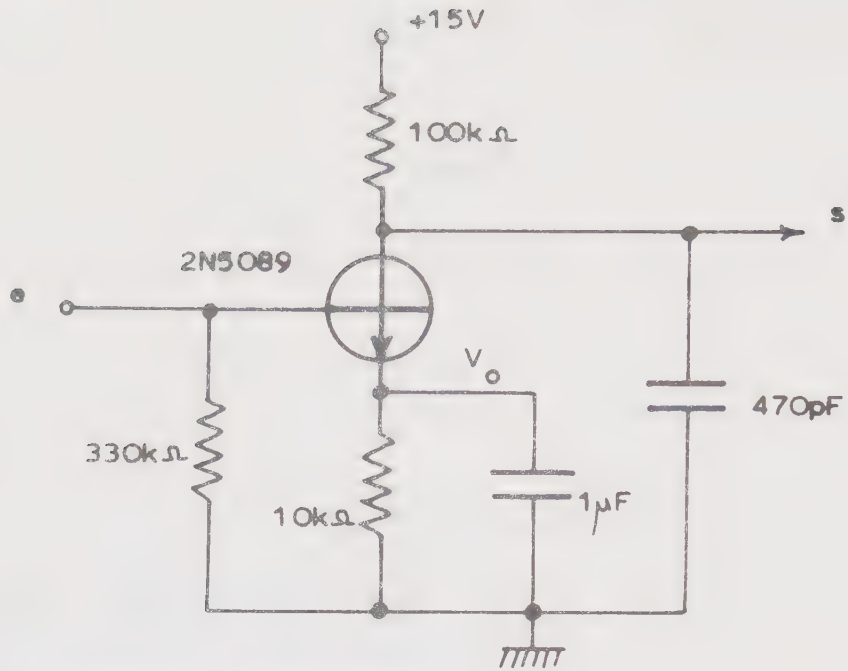
Le détecteur de la voie continue (figure 28 b) dispose, dans l'émetteur du transistor, d'un élément fournissant une tension continue V_0 indépendants de la fréquence, et équivalent à un système à constante de temps infinie. Le dispositif peut ainsi transmettre les modulations depuis la fréquence nulle.

2.4.3. Etude du filtrage du glottogramme

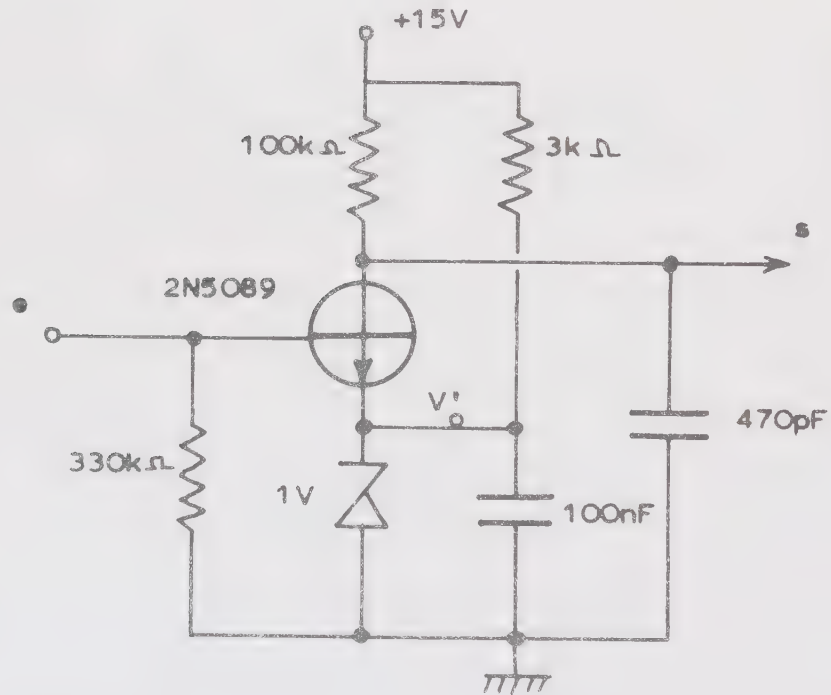
Le signal issu des détecteurs de modulation d'amplitude est entaché de composantes parasites qu'il est possible d'atténuer à l'aide d'un filtrage permettant de ne conserver que les signaux aux fréquences utiles (0 à 600 Hz pour la voie continue, 80 à 600 Hz pour la voie alternative), sans déformation de leur forme.

Les bruits à fréquence élevée provenant des composants et des résidus de porteuse peuvent être réduits par un filtrage passe-bas dont la fréquence de coupure doit permettre de transmettre un nombre suffisant d'harmoniques de la fréquence laryngée la plus élevée (600 Hz) pour conserver la forme du signal avec une bonne précision.

Nous avons vu, au (§ 2.3) que le spectre du glottogramme décroît de 12 dB par octave, c'est-à-dire que l'amplitude de l'harmonique de rang n est égale à celle du fondamental



a) Voie alternative;



b) Voie continue;

Figure 28 - DETAILS DES DETECTEURS DU GLOTTOGRAPHE .

divisée par n^2 . Dans ces conditions, la composante d'ordre 8 vaut 1,5 % de la fondamentale. Compte tenu des fluctuations d'amplitude dues aux bruits, une fréquence de coupure de 5000 Hz se situant immédiatement après l'harmonique 8 de la plus haute fréquence laryngée est acceptable. Le rapport signal sur bruit étant de l'ordre de 30 dB, la précision de l'amplitude du glottogramme est inférieure à 5 %. Après plusieurs essais, nous avons choisi un filtre de Legendre passe-bas d'ordre 3 dont la transmittance est :

$$F(p) = \frac{1}{(1 + 1,612 p) (1 + 0,742 p + 1,074 p^2)} .$$

La détection sur la voie alternative ne peut atténuer efficacement, par son filtrage sommaire, les signaux très basse fréquence indésirables dus aux mouvements lents du larynx. Un filtrage passe-haut supplémentaire est nécessaire. Le choix de la fréquence de coupure (f_c), voisine des premiers harmoniques des fréquences basses des émissions laryngées (80 Hz) et de l'ordre n de ce filtre, doit tenir compte du déphasage variable en fonction de la fréquence que peut entraîner, surtout au voisinage de f_c , une déformation du signal parfois gênante. Pour qu'un réseau électrique transmette un signal sans déformation, il suffit qu'il fasse subir un retard τ constant à toutes ses composantes harmoniques. Pour une pulsation ω , ce retard, appelé temps de propagation de groupe, se traduira par un déphasage $\phi = \omega \tau$. L'expérimentation de quelques filtres étudiés par Bildstein [3] nous a conduit à choisir un filtre passe-haut de Legendre d'ordre 3, dont les courbes d'amplitude et de temps de propagation de groupe en fonction de la fréquence (figure 29) permettent de fixer la fréquence de coupure à 30 Hz. Sa transmittance est :

$$F(p) = \frac{p}{1 + 1,612 p} \cdot \frac{p^2}{1 + 0,742 p + 1,074 p^2} .$$

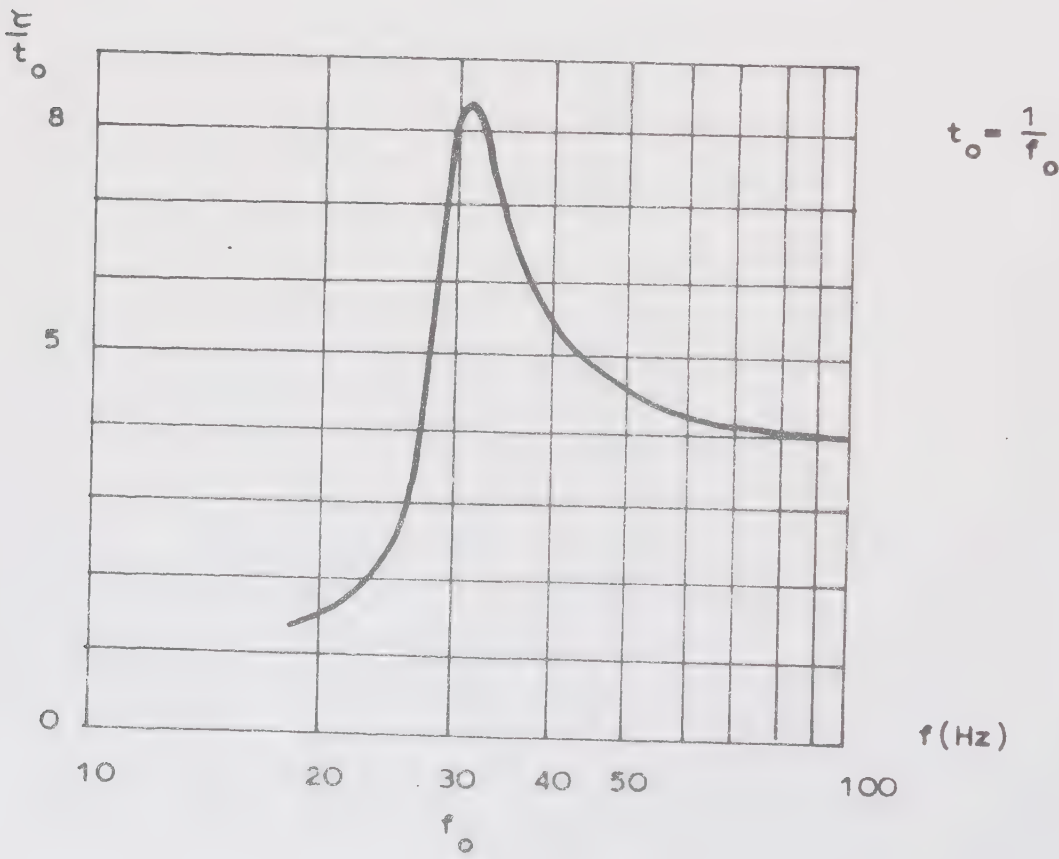
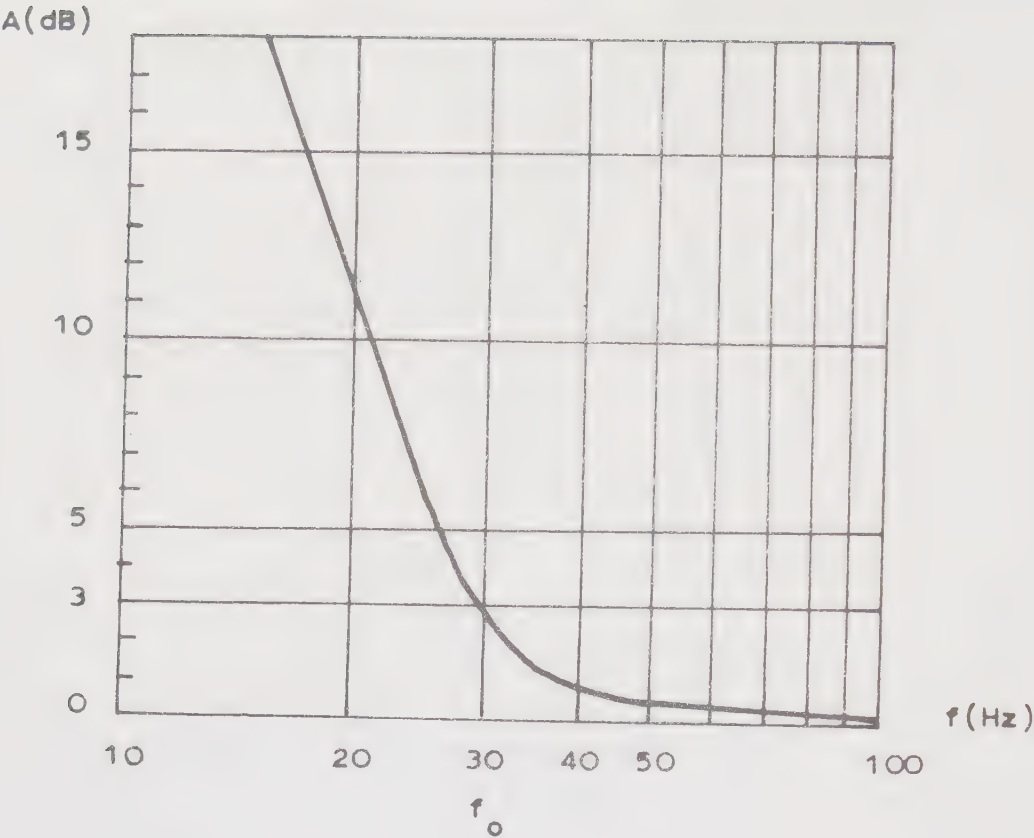


Figure 29 - COURBES D'ATTENUATION ET DE PROPAGATION DE GROUPE DU FILTRE PASSE HAUT .

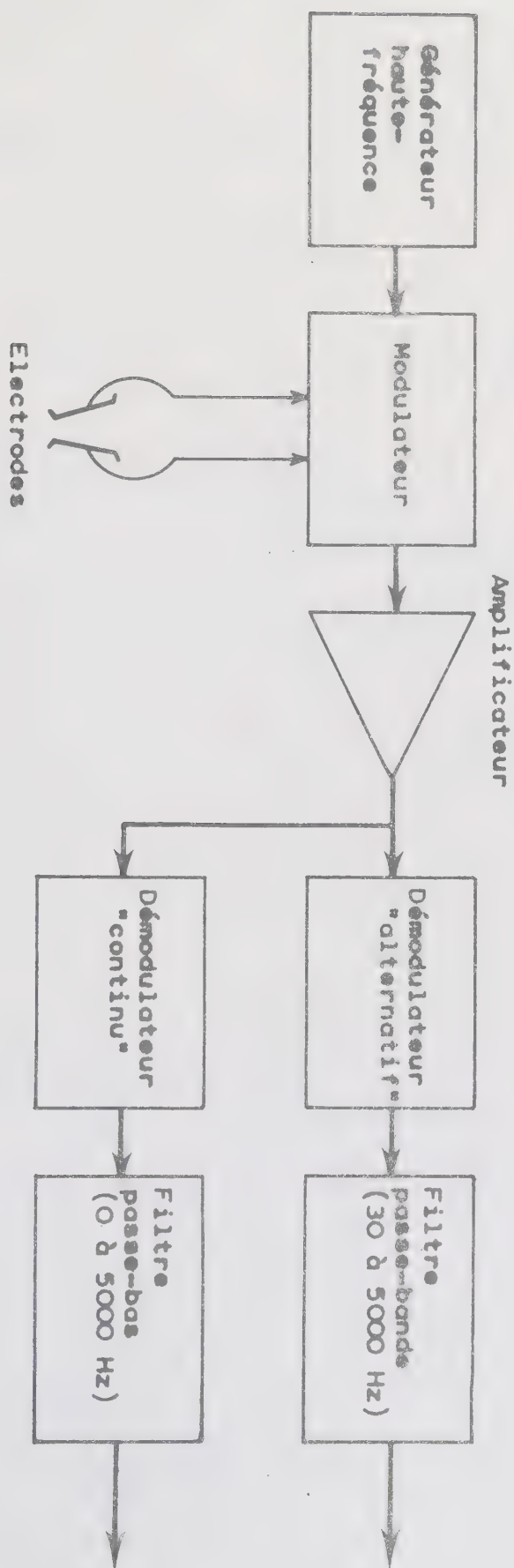


Figure 30 - SCHEMA SYNOPTIQUE DE L'ELECTROGLOTOGRAPHE .

Le sonéma synoptique simplifié de la figure 30 montre les liaisons entre les différents circuits composant le glottographe.

2.4.4. Notice technique de l'électroglottographe

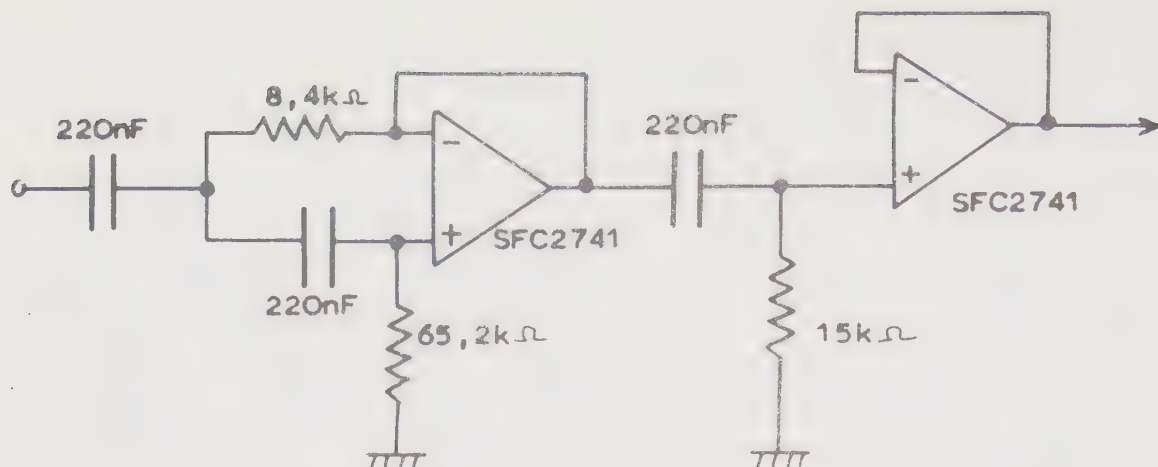
2.4.4.1. Détails de construction

Les schémas de réalisation de l'électroglottographe sont donnés aux figures 31 a, b, c.

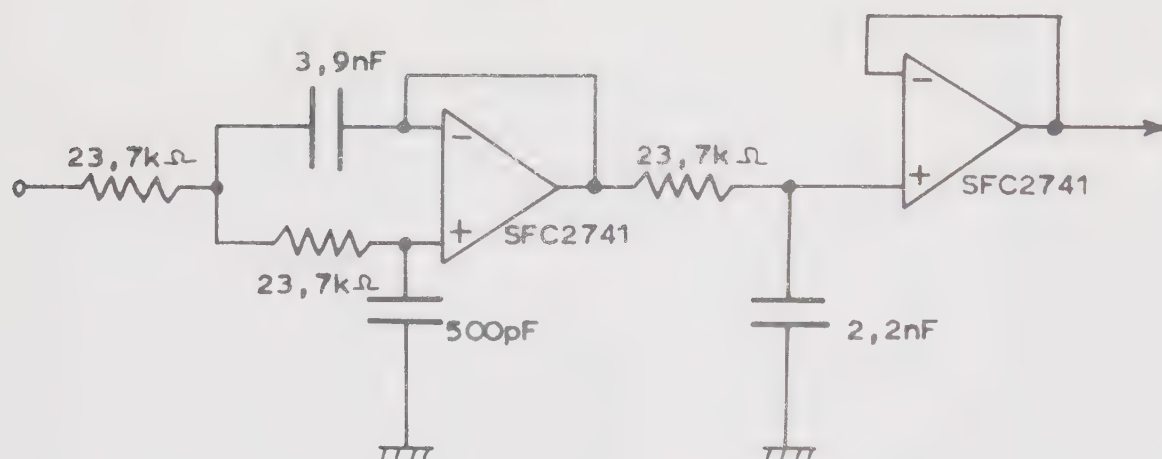
La résistance R du modulateur, en série avec le primaire du transformateur, est constituée d'un potentiomètre P de 10 k Ω en série avec une résistance de 1 k Ω . Il est nécessaire de pouvoir ajuster cette résistance en fonction du module de l'impédance du larynx du locuteur, pour attaquer les détecteurs sous une tension de porteuse du même ordre de grandeur dans chaque cas ; ceci permet de faire constamment travailler les transistors de détection dans la même région de la caractéristique et d'obtenir ainsi des résultats comparables à chaque expérience.

On remarque, sur la voie alternative, un amplificateur qui précède les filtres passe-haut et passe-bas : il fixe le niveau de sortie du glottogramme qui est de 5 V en moyenne.

Sur la voie continue, le circuit qui suit le détecteur sert à soustraire la composante continue induite par la tension collecteur-masse du transistor de détection au repos (sujet inactif); l'équilibrage se fait automatiquement lorsque l'on ajuste la résistance R, de façon à représenter le module de l'impédance (Z) du larynx au repos par une tension quasiment nulle. Cette disposition évite la saturation éventuelle du filtre passe-haut par de trop fortes variations de Z.



Filtre passe-haut
 $f_c = 30\text{Hz}$



Filtre passe-bas
 $f_c = 5000\text{Hz}$

Figure 31 b - FILTRES DE SORTIE .

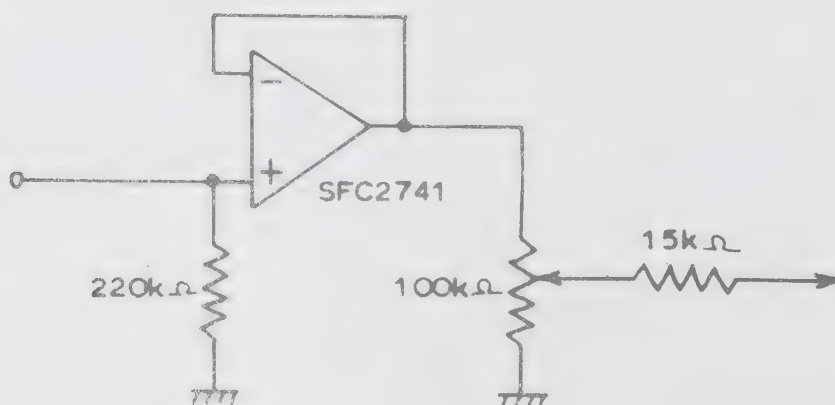


Figure 31 c - CIRCUIT D'ADAPTION A UN ENREGISTREUR
 MAGNETIQUE .

La précision sur les éléments passifs constituant les filtres passe-haut et passe-bas doit être d'environ 2 % pour éviter les ajustements à la construction.

Le collier porte-électrodes est constitué d'une sangle élastique à fermeture réglable et supporte deux électrodes rectangulaires en cuivre doré de 40 x 20 mm; la liaison du collier au glottographe s'effectue à l'aide d'un câble blindé de 80 cm de long, terminé par un connecteur de type DIN à 3 broches.

L'alimentation de l'électroglottographe nécessite deux tensions symétriques (+ 15 V et - 15 V) stabilisées à 0,1 % ; elles sont obtenues à l'aide de deux régulateurs intégrés précédés d'un système classique de redressement et de filtrage. La consommation totale de l'ensemble est inférieure à 200 mA.

Les dimensions du coffret (105 x 160 x 210 mm) font de l'électroglottographe un appareil facile à loger dans une chaîne d'analyse de la parole.

2.4.4.2. Bornes d'accès et commandes

* Bornes d'entrée-sortie :

- embase DIN à trois broches recevant le connecteur du collier d'électrodes ; la tension H.F. aux bornes des électrodes est de 1 V crête à crête;
- borne BNC pour la sortie "alternative" (niveau de sortie : 5 V ; impédance de sortie : 600 Ω ; bande passante : 30 Hz à 5000 Hz à -3 dB);
- borne BNC pour la sortie "continue" (niveau de sortie : 5 V ; impédance de sortie : 600 Ω ;

bande passante : 0 à 5000 Hz à -3 dB) ;

- borne BNC servant au raccordement d'un enregistreur magnétique (niveau réglable de 0 à 5 V ; impédance de sortie 600 Ω ; bande passante : 30 à 5000 Hz à -3 dB).

* Commandes accessibles de l'extérieur

- interrupteur de mise sous tension ;
- potentiomètre d'équilibrage de l'impédance du larynx ;
- atténuateur de la sortie pour enregistrement magnétique.

2.4.4.3. *Méthode d'utilisation*

Le collier porte-électrodes étant connecté et l'électroglottographe mis sous tension, on pose les électrodes, préalablement enduites d'une couche mince de pâte conductrice, de part et d'autre du cartilage thyroïde au niveau de la glotte et on ferme le collier ; il faut ensuite procéder à l'équilibrage de l'impédance du larynx à l'aide du potentiomètre correspondant de façon à amener l'aiguille du vu-mètre au centre ; l'appareil est alors prêt à fonctionner.

Les trois sorties peuvent être utilisées simultanément :

- les sorties "alternative" et "continue" nécessitent un système d'enregistrement graphique rapide tel que les enregistreurs à jet d'encre ou à éclairage ultra-violet ;

- la sortie "enregistrement magnétique" permet d'utiliser un grand nombre d'appareils grâce à son circuit adaptateur et à son atténuation ajustable.

C H A P I T R E 3

DISPOSITIF DE MESURE DE FREQUENCE INSTANTANEE

Nous disposons, à la sortie de l'électroglottographe, d'un signal représentant directement le mouvement des cordes vocales lors de l'émission des sons voisés par le locuteur. La mise en forme de ce signal (c'est-à-dire la production d'impulsions brèves aux instants d'apparition des impulsions laryngées) rendra possible la mesure de la fréquence mélodique instantanée.

Les variations des durées des cycles glottiques successifs interdisent la mesure de fréquence par simple comptage des impulsions laryngées pendant un temps fixe. L'appareil étudié devra nécessairement déterminer la valeur de chaque intervalle de temps séparant l'émission de deux impulsions laryngées. Dans la suite de l'étude, nous caractériserons indifféremment ce type de mesure par les termes de "mesure d'intervalles de temps successifs" ou de "mesure de fréquence instantanée".

Après avoir traité de la mise en forme du signal, ce chapitre expose les contraintes imposées au système et présente les méthodes de mesure connues dont la discussion débouche sur le choix de la solution. Celle-ci nous a amené à réaliser un dispositif original de mesure logarithmique directe d'intervalles de temps successifs, dont les extensions et les transformations possibles ont justifié une étude s'écartant quelque peu de l'analyse de mélodie.

3.1. MISE EN FORME DU GLOTTOGRAMME

L'entrée d'un dispositif de mesure d'intervalles de temps nécessite un traitement préalable du signal analysé appelé "mise en forme", qui consiste à détecter et à

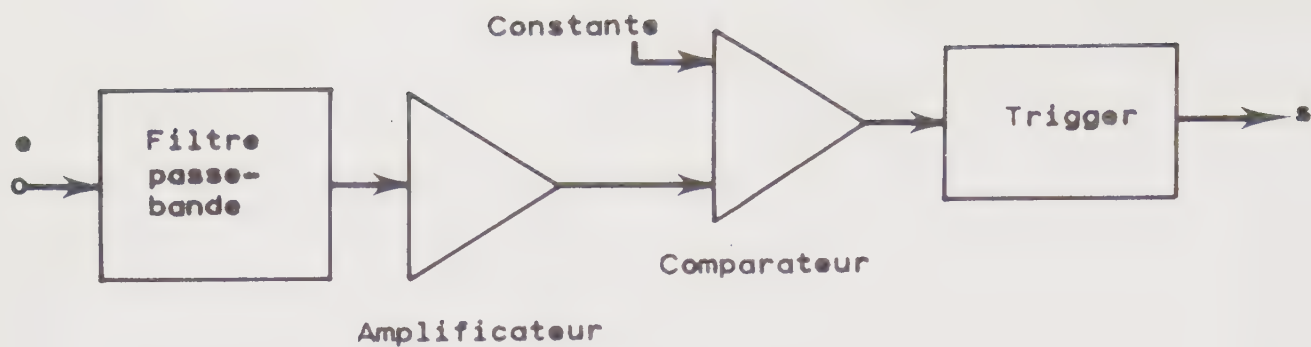
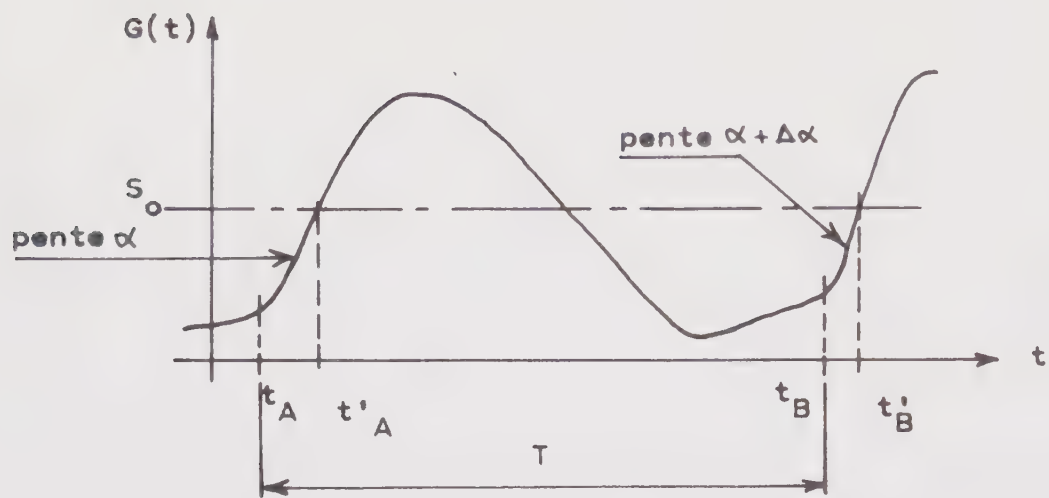
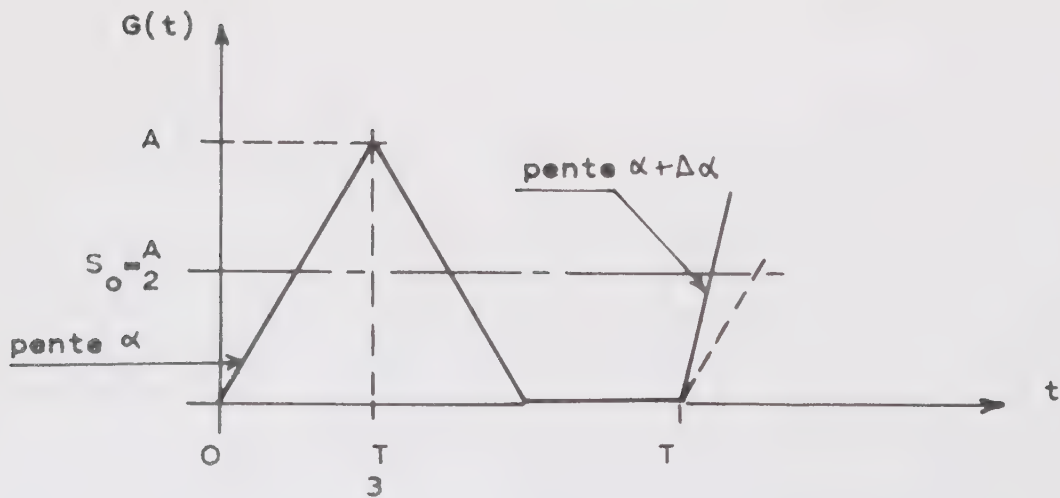


Figure 32 - SCHEMA DE PRINCIPE DU CIRCUIT DE MISE EN FORME .



a) Effet de la variation de pente du glottogramme



b) Glottogramme idéalisé

Figure 33 - MISE EN FORME DU GLOTTOGRAMME .

matérialiser avec précision par des impulsions brèves, les instants de début et de fin de la durée mesurée.

La figure 32 rappelle le schéma de principe d'un circuit de mise en forme faisant apparaître ses différentes fonctions :

- filtrage permettant d'isoler le domaine de fréquences à mesurer
- amplification
- comparaison à un seuil fixe fournissant un changement brusque de niveau à chaque franchissement du seuil
- adaptation des niveaux de sortie du comparateur au circuit de mesure et génération des impulsions.

Le problème que pose la mise en forme d'un signal non sinusoïdal et de forme variable comme le glottogramme est la détermination précise des instants du début et de la fin de la pseudo-période.

La figure 33 a montre, sur un glottogramme usuel, les deux instants successifs t_A et t_B d'ouverture de la glotte, dont l'écart ($t_B - t_A$) est exactement la durée (T) du cycle glottique. Or, il est pratiquement impossible d'atteindre ces points avec le comparateur dont l'indication manquerait de précision avec la proximité du palier quasiment horizontal correspondant à l'accollement des cordes vocales. Il est donc nécessaire de prévoir un seuil de comparaison (S_0) situé dans les régions à pente positive, ce qui fait apparaître une nouvelle possibilité d'erreur, mais moins importante, due à la variation de pente du début du glottogramme, d'une impulsion laryngée à la suivante, provenant d'un changement d'amplitude ou de forme. L'erreur (ΔT) peut être rendue négligeable en amplifiant fortement le signal avant de l'appliquer au comparateur : ΔT est divisée par le coefficient d'amplification. A titre d'exemple, en considérant le signal

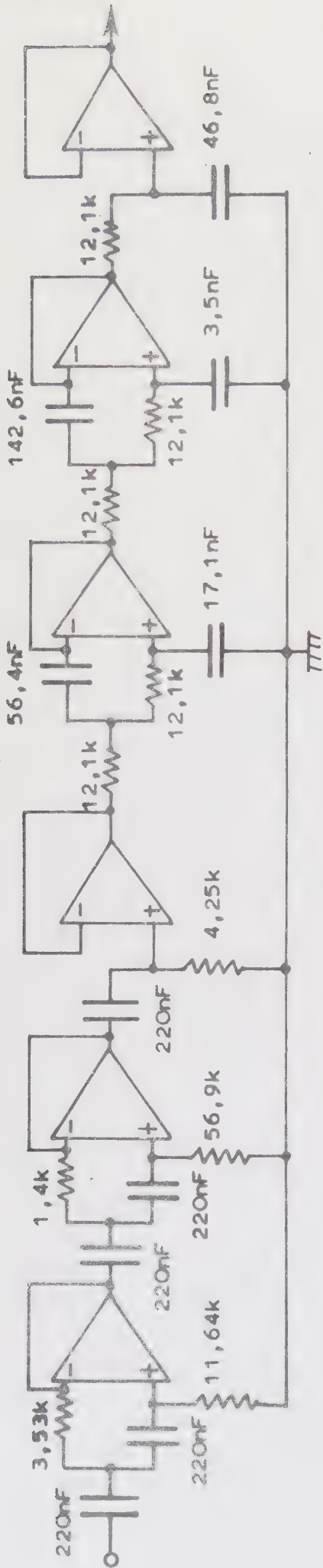


Figure 34 a) Filtre passe-bande d'ordre 5

Amplificateurs
opérationnels : SFC2741

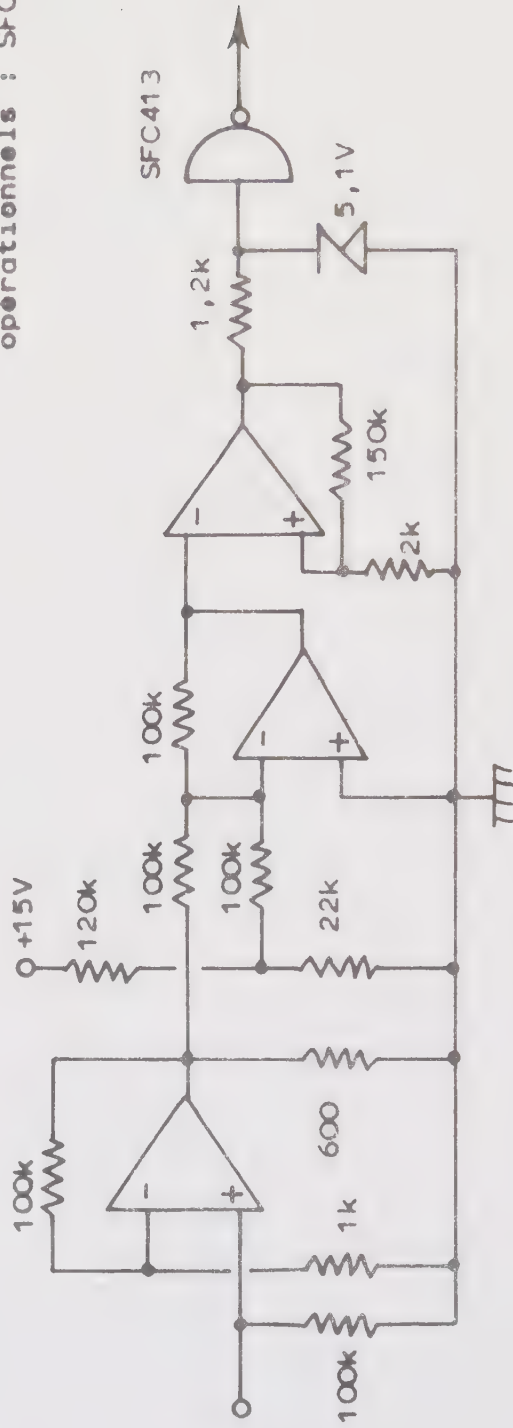


Figure 34 b) Schéma de l'ensemble de mise en forme .

idéalisé de la figure 33 b, avec une variation de pente de 75 % (ce qui est énorme et jamais atteint), l'erreur sur la détection de T est :

$$\frac{\Delta T}{T} = \frac{S_0}{\alpha T} \frac{\frac{\Delta \alpha}{\alpha}}{1 + \frac{\Delta \alpha}{\alpha}} = 0,1$$

(α est la pente initiale et $\Delta \alpha$ la variation de pente au début de la période suivante).

En amplifiant le signal par 100, l'erreur relative sur T devient $\frac{\Delta T}{T} = 10^{-3}$.

Le problème des tensions de bruit superposées au signal utile devient alors critique, car elles sont également amplifiées. Il est donc nécessaire de disposer, en amont du dispositif, d'un filtrage efficace permettant d'éliminer aussi bien les bruits à fréquence élevée que les parasites basse-fréquence comme la modulation accidentelle par le secteur à 50 Hz ou par des tensions résiduelles dues aux mouvements d'ensemble du larynx. C'est la raison pour laquelle nous avons placé, à l'entrée du circuit de mise en forme un filtre passe-bande d'ordre 5 à réponse de Legendre ne laissant passer que les fréquences comprises entre 80 et 600 Hz qui constituent le domaine des émissions laryngées.

La variation non linéaire de phase au voisinage des fréquences de coupure n'est pas gênante, puisque nous ne nous intéressons, dans ce cas, qu'à la période de la première composante harmonique du phénomène, qui est transmise intégralement, même si elle subit un léger retard.

Le schéma de réalisation du circuit de mise en forme est donné à la figure 34.

3.2. CARACTERISTIQUES IMPOSEES A L'APPAREILLAGE DE MESURE

Les spécifications générales concernant les analyseurs de mélodie (§ 1.4.1) peuvent se résumer, dans le cas du système de mesure à :

- fiabilité et précision voisine de 2 % ;
- restitution correcte des variations brusques de hauteur ;
- fonctionnement en temps réel.

Les deux dernières conditions sont satisfaites par le fait que l'appareil devra mesurer la durée de chaque cycle glottique.

La précision de mesure dépend essentiellement de la technologie adoptée.

Nous nous sommes imposé, en plus des caractéristiques précédentes, de fournir le résultat de la mesure sous la forme logarithmique qui est particulièrement apte à représenter des intervalles de fréquence. Cette présentation, couramment utilisée par les musiciens, offre un intérêt pour les phonéticiens dans l'étude de l'intonation, par exemple, et surtout pour caractériser les voyelles des langues à tons. De plus, elle est en rapport direct avec la sensation auditive produite par des variations de hauteur selon la loi de Weber-Fechner, la sensation de hauteur (comme celle d'intensité) est proportionnelle au logarithme de l'excitation (ici la fréquence). Il convient cependant de noter que cette loi ne s'applique pas intégralement à toute la bande des fréquences audibles, et qu'elle s'écarte légèrement de la réalité aux extrêmes. Cependant, des mesures de sensibilité différentielle de hauteur caractérisant la plus petite variation de fréquence perceptible, en fonction de la fréquence, ont permis de

vérifier sa constance à $1/10$ de ton près (soit $1/60$ d'octave) ; cet écart est négligeable.

Le domaine de mesure, incluant les fréquences des émissions laryngées en parole normale pour tous les types de locuteurs, doit s'étendre sur une décade à partir de 70 Hz.

La sortie doit se faire sous forme analogique pour une visualisation immédiate, et numérique pour un dépouillement différé des résultats ;

3.3. LES METHODES DE MESURE DE FREQUENCE INSTANTANEE

3.3.1. Méthodes analogiques

Les méthodes analogiques de mesure d'intervalles de temps successifs réalisent une conversion temps-amplitude.

Leur principe général (figure 35) est l'intégration d'une constante, initialisée par l'impulsion de début de mesure et interrompue par l'impulsion finale. L'amplitude de la sortie de l'intégrateur, à cet instant, est proportionnelle au temps écoulé entre les deux impulsions du signal d'entrée. La conversion logarithmique éventuelle de la mesure est généralement assurée par un amplificateur à caractéristique de transfert d'amplitudes adéquate [22], [6] .

Il existe des dispositifs plus élaborés comme la conversion fréquence-amplitude par décharge quasi hyperbolique de condensateurs [49] ou la conversion logarithmique directe fréquence amplitude [23] .

L'expérience montre que les systèmes de mesure analogiques servant à analyser des phénomènes rapides, comme c'est le cas ici, n'ont pas une précision et

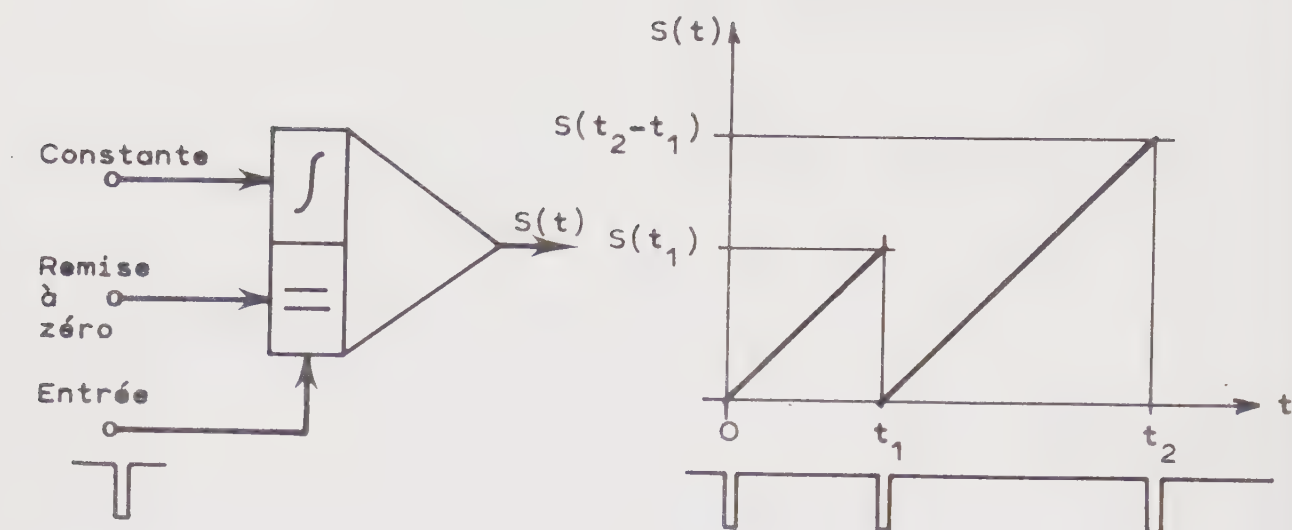


Figure 35 - MESURE ANALOGIQUE D'INTERVALLES DE TEMPS .

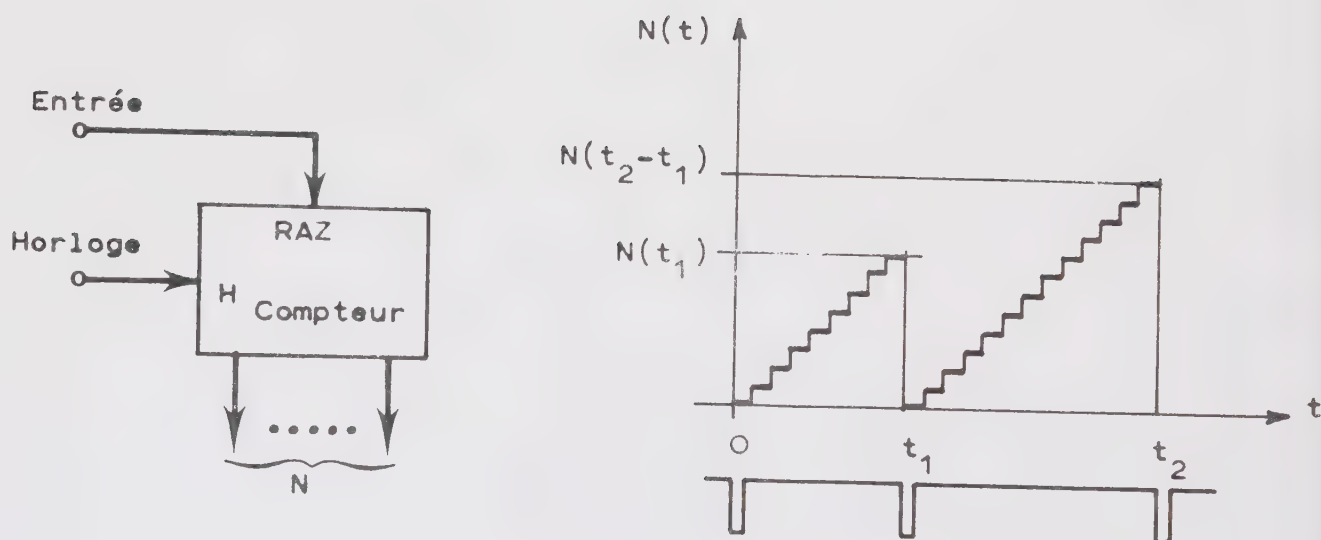


Figure 36 - MESURE DIGITALE D'INTERVALLES DE TEMPS .

une fidélité satisfaisante, du moins pour un prix de revient modéré.

Les techniques numériques, rendues compétitives par la mise au point de composants de plus en plus rapides et miniaturisés, tendent à supplanter les méthodes analogiques dans bien des domaines de mesure.

3.3.2. Méthodes numériques

Le principe de mesure d'intervalles de temps par une méthode numérique (figure 36) consiste, à l'image des systèmes analogiques, à compter les impulsions d'une horloge entre deux impulsions du signal d'entrée. Le contenu du compteur, en fin de période, est proportionnel au temps séparant ces deux impulsions.

Les principaux avantages des composants digitaux sont la fiabilité des ensembles qu'ils composent et la précision dans les fonctions réalisées.

La précision, en mesure de temps, est pratiquement fixée par la valeur et la stabilité de la période d'horloge. L'état d'avancement actuel de ces techniques permet de disposer de cellules de comptage admettant en entrée des fréquences supérieures à 300 MHz.

Ce sont ces raisons qui nous ont conduit à choisir une technologie digitale pour concevoir le module de mesure de fréquence instantanée de l'analyseur de mélodie.

3.4. DISPOSITIF DIGITAL DE MESURE LOGARITHMIQUE DE FREQUENCE INSTANTANEE

3.4.1. Principe général

La plage de mesure d'une décade peut être quantifiée en environ 80 intervalles élémentaires d'un quart de ton, c'est-à-dire d'un vingt-quatrième d'octave puisque'un ton est un sixième d'octave, délimités par les fréquences $F_0, F_1, \dots, F_{80}$ (numérotées à partir de la fréquence maximale F_0). Nous verrons ultérieurement que le choix du quart de ton comme pas de mesure est compatible avec la précision demandée.

Les intervalles élémentaires correspondent à une variation relative de fréquence constante entre deux valeurs discrètes de F .

$$a = 2^{\frac{1}{24}}$$

La valeur de F de rang k a pour expression :

$$F_k = F_0 a^{-k}$$

Nous dirons que F_k se situe à k intervalles élémentaires de F_0 selon la formule :

$$k = -\log_a \frac{F_k}{F_0}.$$

Les périodes T_k correspondant aux valeurs quantifiées de F forment donc une progression géométrique de raison (a) :

$$T_k = a^k T_0.$$

La mesure consiste à déterminer la fréquence F_k immédiatement supérieure à la fréquence instantanée du signal analysé, ou, ce qui est identique, la

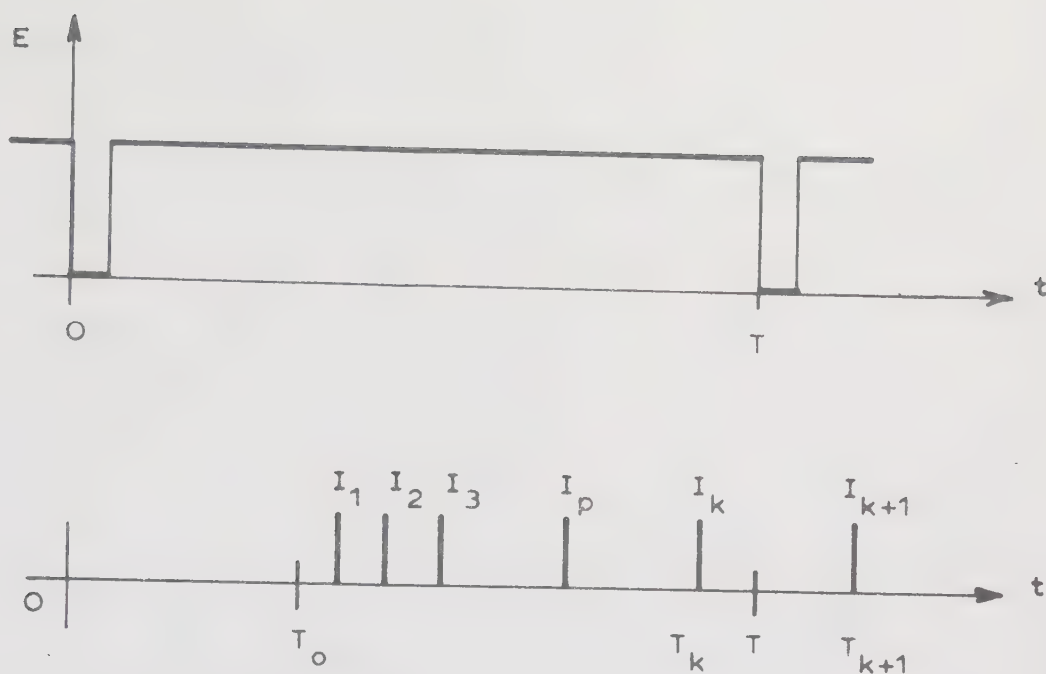


Figure 37 - PRINCIPE GENERAL DU DISPOSITIF DE MESURE DE FREQUENCE INSTANTANNEE .

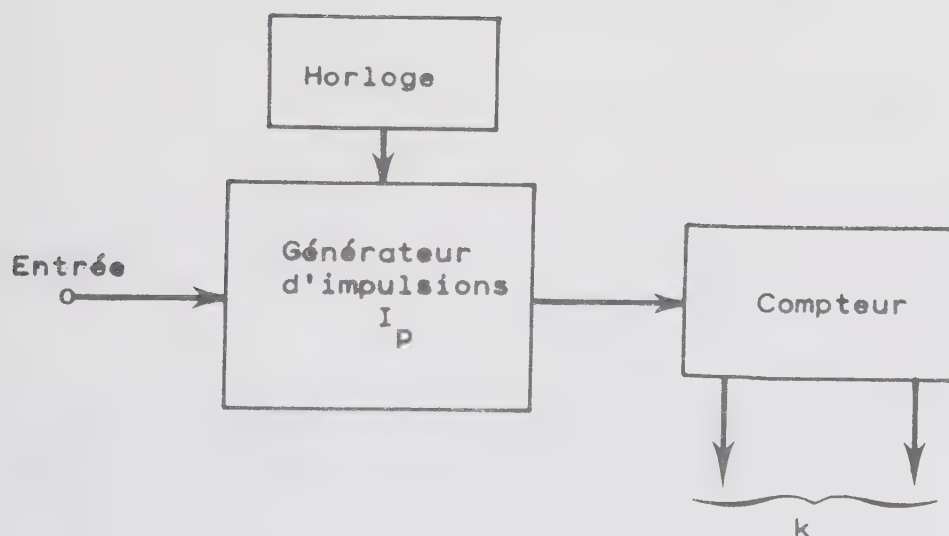


Figure 38 - SCHEMA FONCTIONNEL DU DISPOSITIF DE MESURE DE FREQUENCE INSTANTANNEE .

On peut remédier à cet inconvénient en diminuant la période de référence (T_0) d'un demi-pas, ce qui permet une mesure centrée et une erreur inférieure à $\pm 0,5$ (c'est-à-dire à $1/8$ de ton) (figure 39).
On a alors : $T_k = a^{(k - \frac{1}{2})} T_0$.

L'erreur relative sur T , due à la quantification, est :

$$\left| \frac{\Delta T}{T} \right|_1 = \left| \frac{\Delta(a^x)}{a^x} \right| = |\Delta x| \ln a < \frac{1}{2} \ln a \text{ car } |\Delta x| < 0,5.$$

Notation : $\ln a$ désigne le logarithme népérien de a .

Le choix des instants T_p au voisinage des valeurs réelles T_p parmi les dates fournies par une horloge à fréquence fixe cause également une erreur qui peut atteindre une demi-période ($t/2$) de l'horloge. On détermine donc la valeur de t en se fixant l'erreur admissible sur la position des impulsions T_p autour des T_p . Cette erreur est maximale pour $T_p = T_0$, et l'erreur relative correspondante sur T est :

$$\left| \frac{\Delta T}{T} \right|_2 = \frac{t}{2T_0} \text{ d'où } t = 2 \left| \frac{\Delta T}{T} \right|_2 T_0$$

L'erreur relative totale sur la valeur mesurée de T est donc telle que :

$$\left| \frac{\Delta T}{T} \right| < \frac{t}{2T_0} + \frac{1}{2} \ln a = \frac{t}{2T_0} + \frac{\ln 2}{48}$$

Dans le cas de notre réalisation, nous obtenons :

- erreur due à la quantification : $\left| \frac{\Delta T}{T} \right|_1 < 1,5 \%$

- erreur due à l'échantillonnage : nous avons fixé $\left| \frac{\Delta T}{T} \right| < 0,5 \%$, ce qui impose une période d'horloge

$$t = \frac{T_0}{100}.$$

3.4.3. Générateur d'impulsions à distribution exponentielle

3.4.3.1. Synthèse globale absolue

Comme nous désirons prélever l'impulsion I_p parmi les impulsions fournies par l'horloge primaire, nous ne pouvons obtenir qu'une valeur approchée multiple entier de la période primaire $T_0/100$:

$$T_p = (100 a^p) \frac{T_0}{100}$$

où $(100 a^p)$ désigne la meilleure approximation entière de $100 a^p$, soit :

$$(100 a^p) = \text{Partie entière} \left(100 a^p + \frac{1}{2} \right)$$

Ce prélèvement peut s'obtenir de la manière suivante :

- les différentes valeurs successives de $(100 a^p)$ sont inscrites dans une table séquentielle S, tandis que les impulsions d'horloge primaire émises depuis le début de la période analysée sont dénombrées par un compteur C ;
- l'impulsion I_p est élaborée lors de la coïncidence des contenus de S et de C, coïncidence détectée par un comparateur ; cette impulsion sert d'une part, à incrémenter le compteur de sortie et d'autre part à faire progresser d'un pas la table séquentielle.

Ce procédé nécessite un compteur C de capacité 1000 et la possibilité de stocker dans S 80 nombres compris entre 100 et 825 soit 80 mots binaires de 10 bits.

P	(100.a ^P) [*]
0	100
1	103
2	106
3	109
4	112
5	115
6	118
7	122
8	126
9	130
10	134
11	138
12	142
13	146
14	150
15	154
16	159
17	164
18	169
19	174
20	179
21	184
22	189
23	194

Figure 40 -
VALEURS DE (100a^P)^{*} .

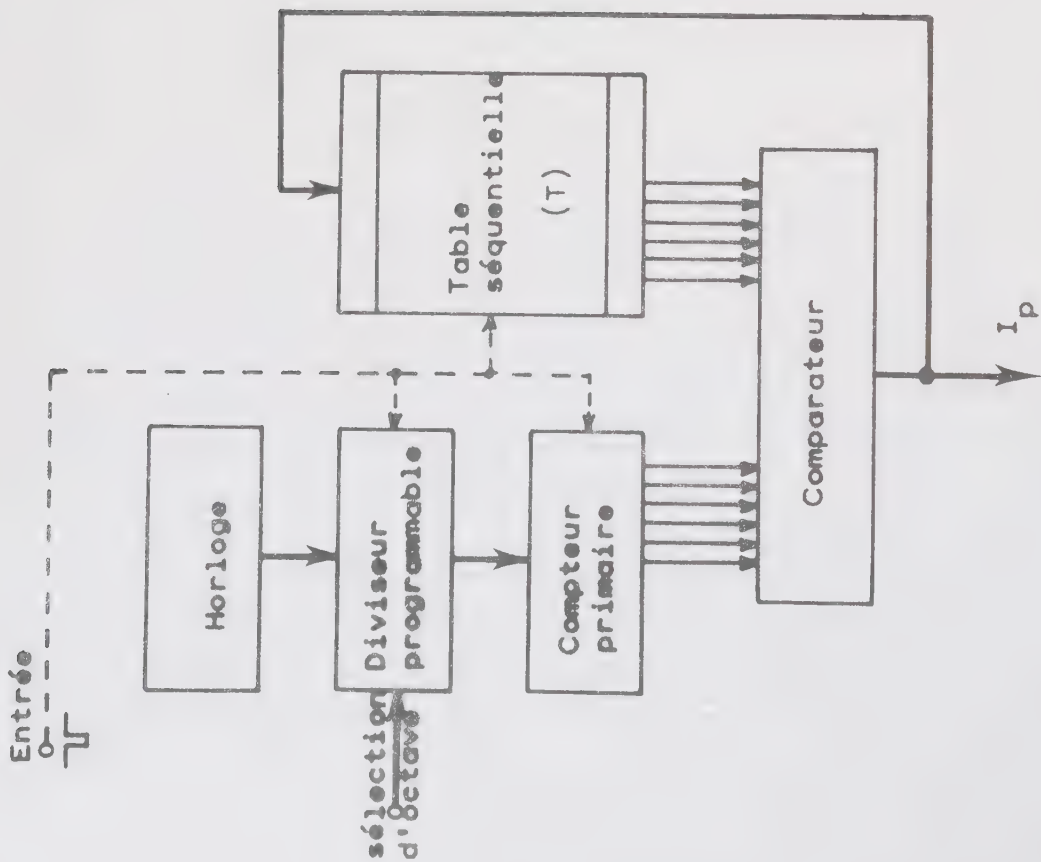


Figure 41- SCHEMA FONCTIONNEL DU GENERATEUR
ABSOLU PAR OCTAVE .

3.4.3.2. Synthèse absolue par octave

Il est heureusement possible de simplifier le système précédent en remarquant que, pour toute valeur de p :

$$T_{p+24} = 2 T_p .$$

Il suffit donc de réaliser un appareil couvrant une octave (14 quarts de ton), le passage à l'octave suivante se faisant simplement en doublant la période d'horloge.

Le chiffre des centaines de $(100 a^p)^*$ étant constamment égal à 1 (figure 40), il suffit de ranger dans S 24 nombres décimaux de 2 chiffres compris entre 00 et 94, soit 24 mots binaires de 7 bits (ou 8 bits en DCB).

Le schéma fonctionnel du dispositif (figure 41) fait apparaître :

- le diviseur de fréquence programmé D , élaborant à partir de l'horloge primaire un signal d'horloge secondaire H_s de période $2^{q-1} \frac{T_0}{100}$ pendant la $q^{i\text{ème}}$ octave; dans notre cas, 4 octaves suffisent pour couvrir la plage de mesure d'une décade;
- la table séquentielle S contenant les 24 valeurs de $100(a^p-1)^*$;
- le compteur primaire C ;
- le comparateur.

P	$(100(a^{P-1}))$	S ₇	S ₆	S ₅	S ₄	S ₃	S ₂	S ₁	S ₀
0	00	0	0	0	0	0	0	0	0
1	03	0	0	0	0	0	0	1	1
2	06	0	0	0	0	0	1	1	0
3	09	0	0	0	0	1	0	0	1
4	12	0	0	0	1	0	0	1	0
5	15	0	0	0	1	0	1	0	1
6	18	0	0	0	1	1	0	0	0
7	22	0	0	1	0	0	0	1	0
8	26	0	0	1	0	0	1	1	0
9	30	0	0	1	1	0	0	0	0
10	34	0	0	1	1	0	1	0	0
11	38	0	0	1	1	1	0	0	0
12	42	0	1	0	0	0	0	1	0
13	46	0	1	0	0	0	1	1	0
14	50	0	1	0	1	0	0	0	0
15	54	0	1	0	1	0	1	0	0
16	59	0	1	0	1	1	0	0	1
17	64	0	1	1	0	0	1	0	0
18	69	0	1	1	0	1	0	0	1
19	74	0	1	1	1	0	1	0	0
20	79	0	1	1	1	1	0	0	1
21	84	1	0	0	0	0	1	0	0
22	89	1	0	0	0	1	0	0	1
23	94	1	0	0	1	0	1	0	0

Figure 42 - TABLE DE CODAGE DES ETATS DU GENERATEUR
ABSOLU PAR OCTAVE

P	Δp	Y ₃	Y ₂	Y ₁	Y ₀	S ₂	S ₁	S ₀
0	3	0	0	0	0	0	1	1
1	3	0	0	0	1	0	1	1
2	3	0	0	1	0	0	1	1
3	3	0	0	1	1	0	1	1
4	3	0	1	0	0	0	1	1
5	3	0	1	0	1	0	1	1
6	4	0	1	1	0	1	0	0
7	4	0	1	1	1	1	0	0
8	4	1	0	0	0	1	0	0
9	4	1	0	0	1	1	0	0
10	4	1	0	1	0	1	0	0
11	4	1	0	1	1	1	0	0
12	4	0	0	0	0	1	0	0
13	4	0	0	0	1	1	0	0
14	4	0	0	1	0	1	0	0
15	5	0	0	1	1	1	0	1
16	5	0	1	0	0	1	0	1
17	5	0	1	0	1	1	0	1
18	5	0	1	1	0	1	0	1
19	5	0	1	1	1	1	0	1
20	5	1	0	0	0	1	0	1
21	5	1	0	0	1	1	0	1
22	5	1	0	1	0	1	0	1
23	6	1	0	1	1	1	1	0

Figure 43 - TABLE DE CODAGE DES ETATS DU

La synthèse de la table séquentielle s'effectue en considérant qu'il s'agit d'un compteur synchrone dont la capacité est le nombre de pas dans l'octave et fonctionnant selon une table de codage très particulière (figure 42) associant au nombre p d'impulsions reçues le code D.C.B de $100(a^p - 1)^*$.

La réalisation de S , que nous avons testée et qui donne entière satisfaction, nécessite l'emploi de 8 bistables JK et d'un ensemble de portes logiques de commande, au total 11 circuits intégrés.

Ce procédé peut cependant être encore amélioré dans le sens de la simplification des circuits.

3.4.3.3. *Synthèse incrémentale par octave*

Les accroissements successifs (arrondis à la valeur entière la plus proche) des nombres $100 a^p$, soit :

$\Delta p = \text{Partie entière} (100 a^{p+1} - 100 a^p + \frac{1}{2})$,
ne prennent que 4 valeurs distinctes (3, 4, 5 et 6) dans toute l'octave (figure 43).

On peut donc n'inscrire dans S que ces 4 chiffres qui apparaissent chacun moins de 10 fois. Le mot binaire de sortie de la table sera constitué de 3 bits S_2 , S_1 et S_0 matérialisés par 3 bistables JK. Les états internes distincts correspondant à des valeurs constantes de Δp pourront être différenciés à l'aide de 4 variables secondaires,

pour lesquelles la structure de la table (figure 43) permet de choisir les 4 bits y_3, y_2, y_1 et y_0 d'un compteur binaire de capacité 12 initialement à 0 et défilant deux fois par octave.

La synthèse de S est alors fort simple et met en oeuvre :

- un compteur de capacité 12 (disponible sous forme d'un circuit intégré) ;
- trois bistables JK associés à quelques portes réalisant les fonctions suivantes :

$$J_0 = S_1 + \bar{y}_3 \bar{y}_2 y_1 \quad J_1 = S_0 y_1 y_3 \quad J_2 = y_0 y_2$$

$$K_0 = y_1 y_3 \quad K_1 = y_0 y_2 \quad K_2 = S_1$$

où J_i et K_i représentent les entrées de commande du bistable numéro i .

Ce mode de travail incrémental de la table séquentielle implique un fonctionnement nouveau pour le compteur C : après un premier comptage jusqu'à 100 (représentant la période de base T_0), le compteur sera remis à zéro à chaque coïncidence de son contenu avec la valeur Δp inscrite dans S, donc remis à zéro par l'impulsion I_p délivrée par le comparateur.

C'est cette solution que nous avons finalement adoptée, car elle ne fait appel qu'à un comparateur à 3 bits et conduit à une synthèse très simple des dates de la table séquentielle.

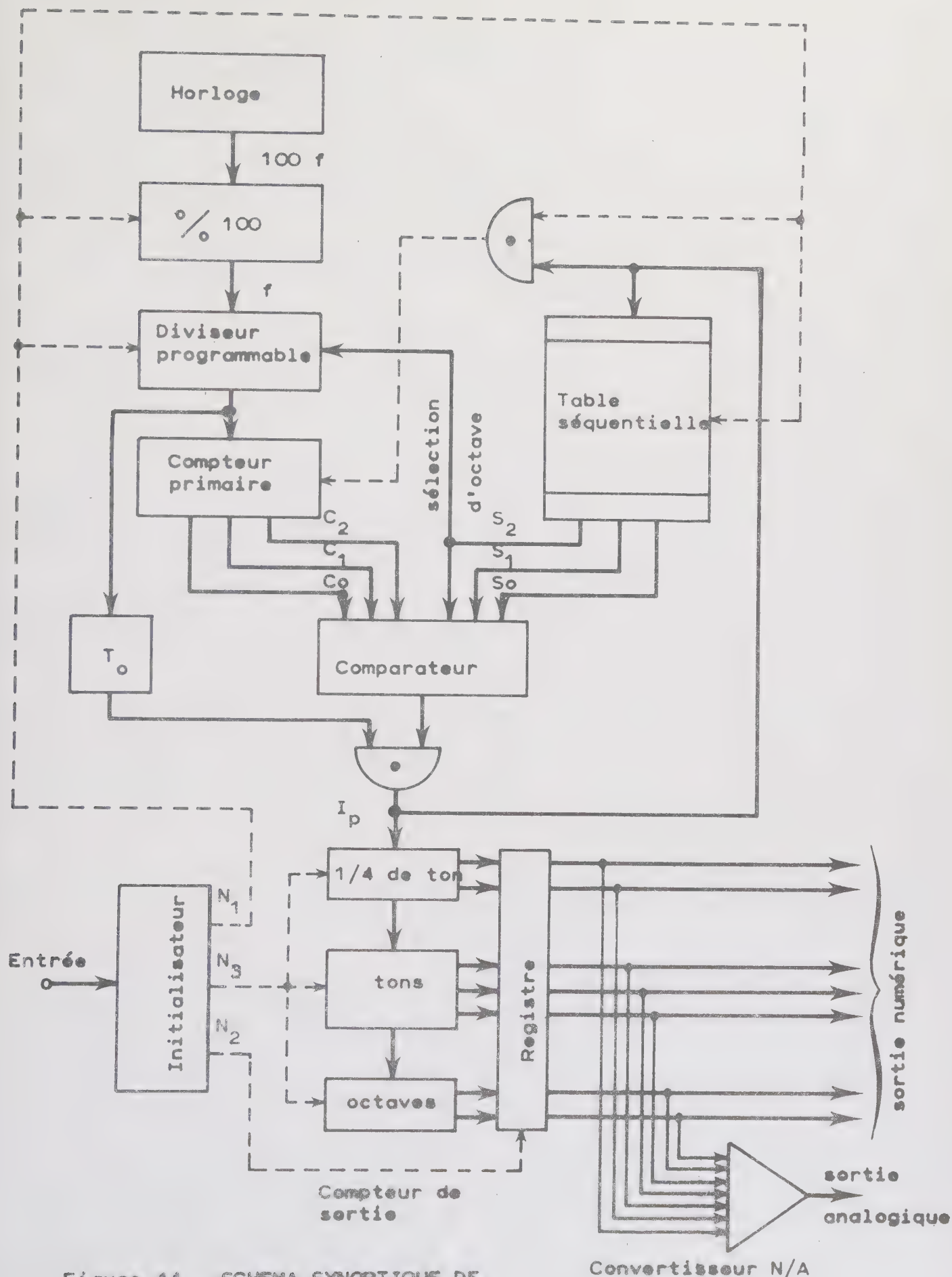


Figure 44 - SCHEMA SYNOPTIQUE DE L'APPAREIL DE MESURE DE FREQUENCE INSTANTANEE .

3.4.4. Réalisation de l'appareil

La figure 44 représente le schéma synoptique détaillé de l'appareil de mesure logarithmique de fréquence instantanée que nous avons réalisé.

Celui-ci se compose de cinq modules principaux :

- l'horloge, incluant la base de temps, le diviseur d'horloge primaire et le dispositif de sélection d'octave ;
- le générateur d'impulsions à distribution exponentielle constitué du compteur primaire (C), de la table séquentielle (S) et du comparateur ;
- le compteur de sortie, ou compteur d'intervalles élémentaires ;
- les dispositifs de présentation de la valeur de la mesure : le registre tampon (sortie numérique) et le convertisseur digital-analogique (sortie analogique) ;
- le système d'initialisation, déclenché par le signal d'entrée.

Les détails de construction de ces modules, repris dans les lignes suivantes, sont illustrés par les schémas de réalisation des figures 45 à 50.

3.4.4.1. *Horloge*

La séquence de mesure est initialisée par une impulsion du signal d'entrée, mais l'horloge interne (ou éventuellement externe) n'est pas déclenchée par ce signal, et il résulte une absence de synchronisation qui, sans créer d'erreur supplémen-

taire, peut apporter pour un signal périodique, une certaine instabilité du contenu k du compteur de sortie représentant la mesure (oscillation entre deux valeurs voisines de k). On remédie à cet inconvénient en élaborant le signal d'horloge primaire de période t par division de fréquence d'une horloge de fréquence nettement supérieure (100 fois par exemple), et en initialisant le diviseur par le signal d'entrée.

Dans notre réalisation, la fréquence maximale à mesurer est de 600 Hz, ce qui impose une horloge primaire de 60 kHz et une horloge générale de 6 MHz. La division par 100 d'une telle fréquence nécessite deux décades de technologie TTL.

En vue de simplifier au maximum la réalisation de l'appareil, nous avons d'abord choisi, pour la base de temps de notre prototype, un multivibrateur construit autour d'un trigger de Schmidt disponible sous forme de circuit intégré et délivrant un signal directement compatible avec la logique du système (figure 45). Cet oscillateur possède une stabilité en fréquence à court terme d'environ 10^{-3} ; elle est suffisante, dans notre cas, compte tenu de la précision demandée. Cependant, comme tout oscillateur à relaxation, sa stabilité à long terme est médiocre, ce qui nécessite un étalonnage fréquent. Pour supprimer cet inconvénient, nous avons remplacé le multivibrateur de la base de temps par un oscillateur harmonique à self et transistor suivi d'un

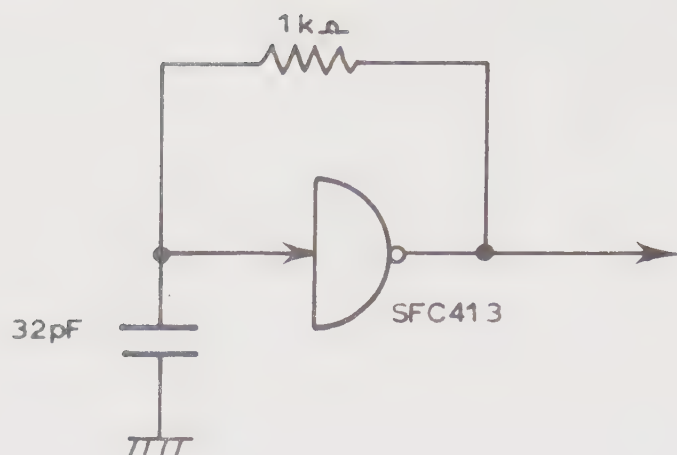


Figure 45 - BASE DE TEMPS A MULTIVIBRATEUR .

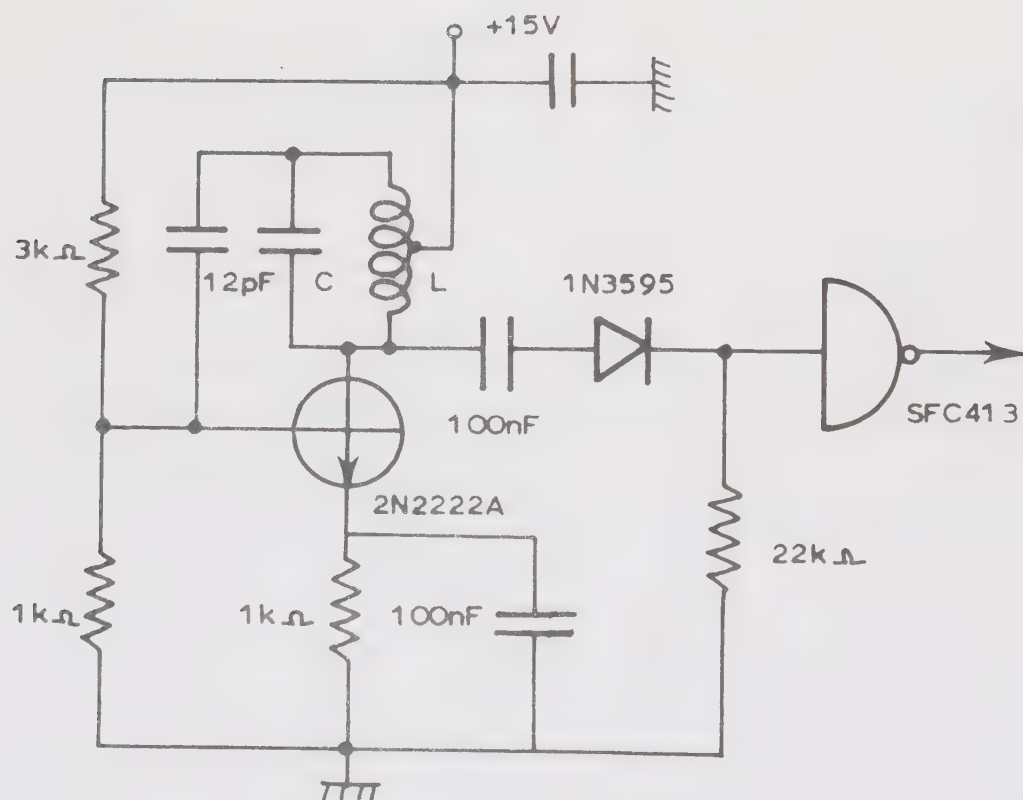


Figure 46 - BASE DE TEMPS A OSCILLATEUR A SELF .

circuit de mise en forme (figure 46) ; cette disposition complique quelque peu la réalisation, mais permet d'obtenir une stabilité en fréquence compatible avec la précision voulue en toutes circonstances, l'utilisation d'un quartz étant superflue.

Le diviseur programmable de sélection d'octave est constitué de 3 bistables JK en cascade, montés en diviseur par deux, et associés à une logique d'aiguillage séquentiel des 4 signaux de période t , $2t$, $4t$ et $8t$ ainsi obtenus vers l'entrée du compteur C. La commande d'aiguillage est réalisée au moyen d'un compteur synchrone de capacité 4 (2 bistables JK) déclenché par le passage de 1 à 0 du bit S_2 de la table séquentielle.

3.4.4.2. *Générateur d'impulsions à distribution exponentielle*

La table séquentielle de type incrémental utilisée, a été décrite au § 3.4.3.3. Elle nécessite un compteur binaire par 12, trois bistables JK et quelques portes formant un ensemble de 5 circuits intégrés.

Le comparateur à 3 bits est constitué d'un seul circuit intégré réalisant la fonction. La coïncidence par paire de toutes ses entrées provoque un changement d'état de sa sortie donnant l'impulsion I_p ; celle-ci incrémente la table séquentielle et remet le compteur C à 0.

Le compteur primaire C doit évoluer une première fois de 0 à 100 au début de la période à mesurer pour matérialiser le

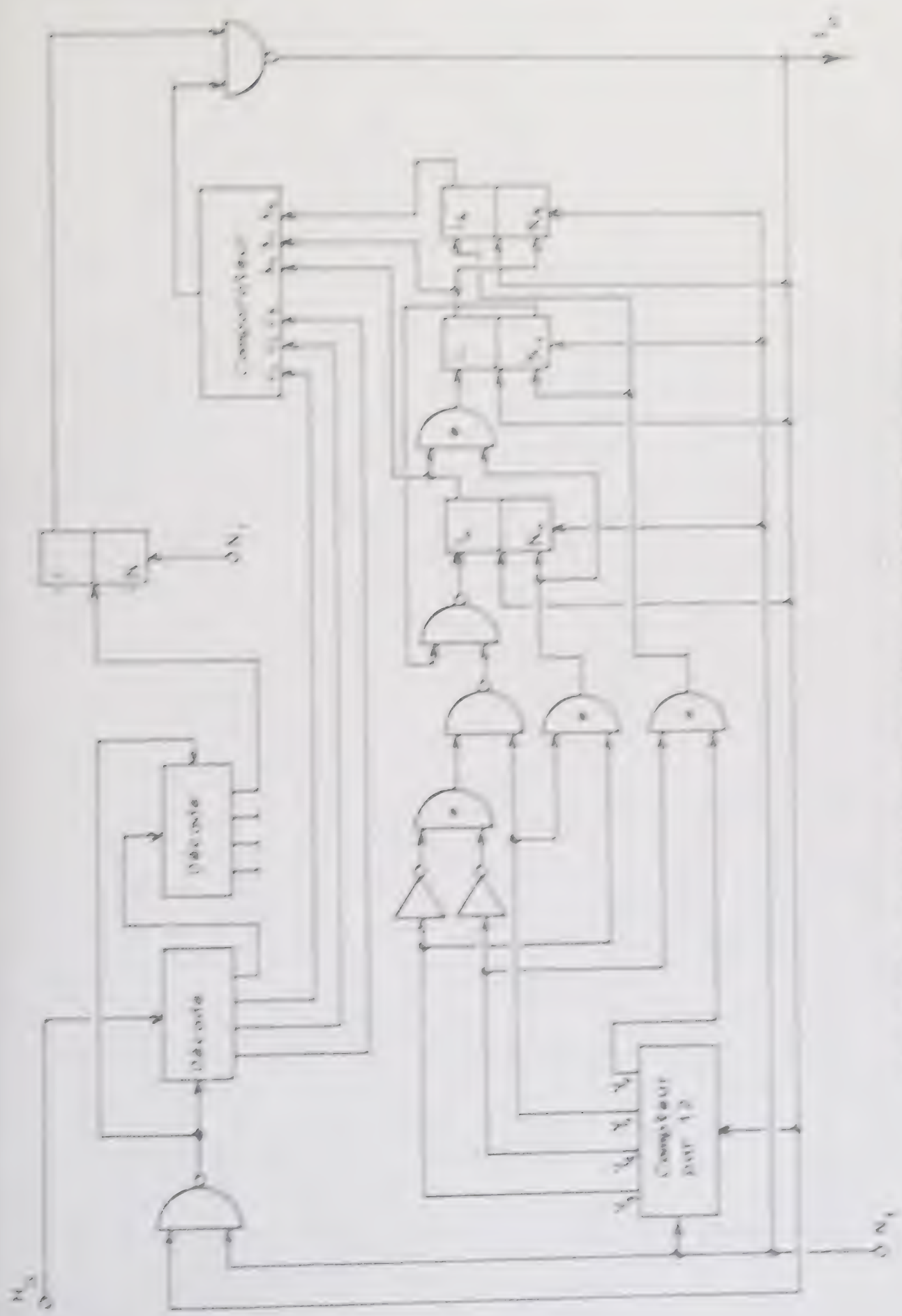


Figure 4.0 - Generation of a 4-bit adder-subtractor circuit

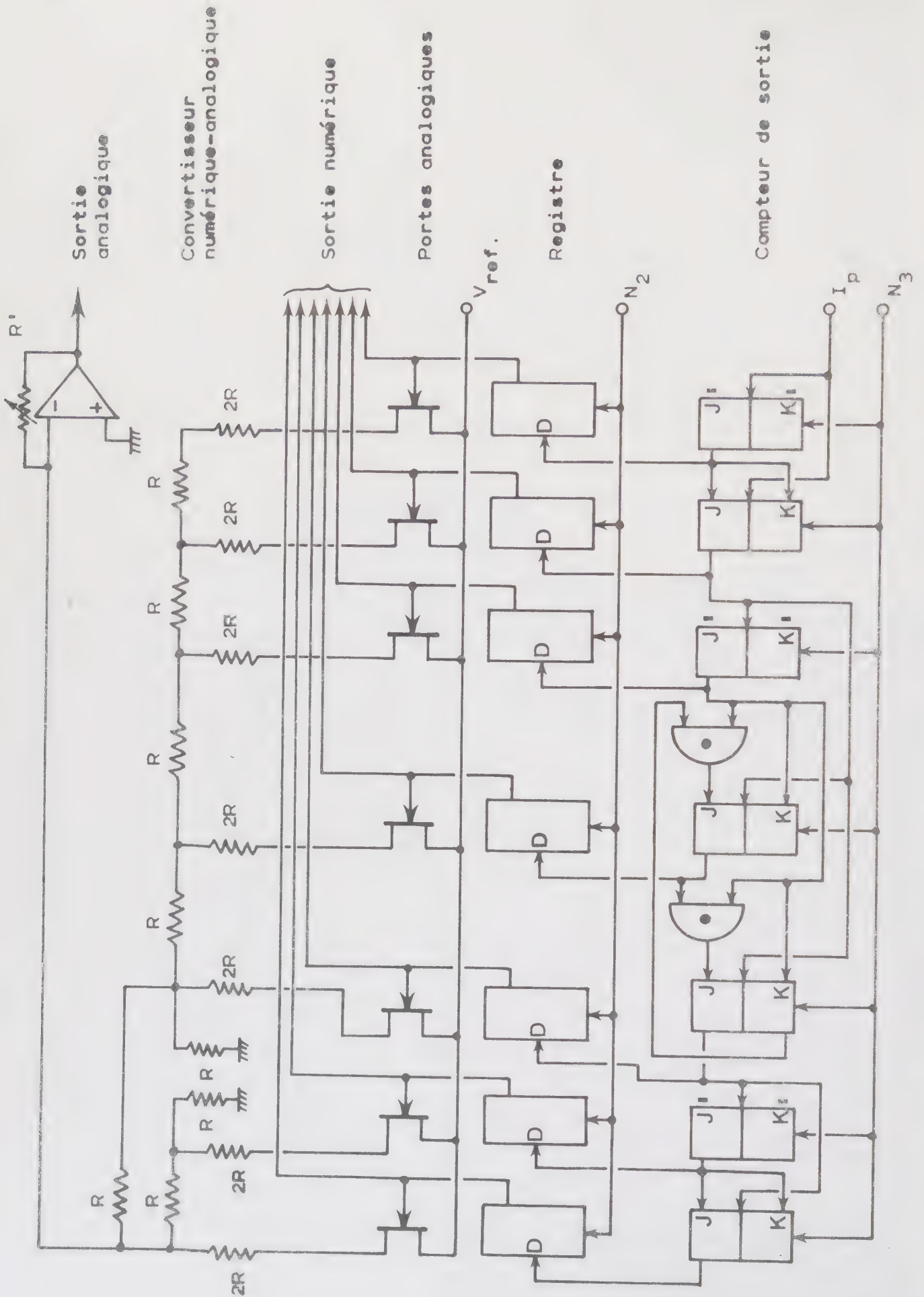


Figure 49 - COMPTEUR DE SORTIE, REGISTRE ET CONVERTISSEUR N/A.

temps de base T_0 ; ce fonctionnement est réalisé en utilisant deux décades en cascade pour C et une bascule déclenchée par le passage à 0 du dernier bit de C ; la sortie de cette bascule bloque, par l'intermédiaire d'une porte, le passage de l'impulsion I_p vers S et la remise à zéro de C tant que le temps T_0 n'est pas écoulé.

3.4.4.3. *Compteur d'intervalles élémentaires*

Le nombre k de fractions de tons représentant la mesure, pourrait être obtenu par un simple compteur décimal ; cependant, tout en conservant la même unité (le quart de ton), il est préférable d'en faire apparaître les multiples employés de manière classique en acoustique musicale :

- le ton valant 4 unités ;
- l'octave valant 6 tons, donc 24 unités.

Ceci nécessite l'utilisation de 3 compteurs binaires placés en cascade (figure 49) :

- un compteur de quarts de ton de capacité 4 constitué de 2 bistables q_0 et q_1 codant le nombre Q de quarts de tons :

$$Q = 2q_1 + q_0 .$$
- un compteur de tons de capacité 6 constitué de 3 bistables t_2 , t_1 et t_0 codant le nombre T de tons :

$$T = 4t_2 + 2t_1 + t_0 .$$
- un compteur d'octaves de capacité 4 constitué de 2 bistables σ_0 et σ_1 codant le nombre O d'octaves :

$$O = 2\sigma_1 + \sigma_0 .$$

En définitive, k se trouve représenté par un nombre de 3 chiffres (OTQ) tel que :

$$k = 24.O + 4.T + Q$$

et le résultat de la mesure s'exprime ainsi :

"la fréquence F est à O octaves, T tons et Q quarts de ton en dessous de la fréquence de référence F_0 ".

La réalisation du compteur de sortie met en oeuvre 7 bistables JK et deux portes ET. Il est à remarquer que le compteur d'octaves peut être confondu avec le compteur de commande du dispositif de sélection d'entrée qui réalise la même fonction.

3.4.4.4. *Registre de sortie et convertisseur digital-analogique*

En fin de comptage, le contenu du compteur est transféré dans un registre tampon constitué de sept bascules de type "D", qui présente et maintient la valeur acquise de k, correspondant à la période écoulée, jusqu'à la fin de la période suivante.

Ce registre est la sortie numérique de l'appareil qui peut être connectée à tout dispositif digital d'enregistrement ou de traitement. Elle est également visualisée au moyen de 3 indicateurs décimaux (O,T,Q) à diodes électroluminescentes ; la mesure d'un signal périodique à fréquence ajustable permet ainsi d'effectuer l'étalonnage en "unités logarithmiques de fréquence du document servant à l'enregistrement des mesures.

La sortie analogique est réalisée par l'addition pondérée des contenus des sept cellules du registre tampon :

$$k = 48 \sigma_1 + 24 \sigma_0 + 16 t_1 + 8 t_1 + 4 t_0 + 2q_1 + q_0 .$$

Ce convertisseur digital-analogique (figure 51) dont les poids ne correspondent pas à un codage usuel, est constitué d'un réseau de résistances (R) égales (2 R étant obtenue par mise en série de deux résistances R) associé à un amplificateur opérationnel.

La valeur de la tension de sortie S_A est :

$$S_A = - \frac{ER'}{R} \left(\frac{q_0}{96} + \frac{q_1}{48} + \frac{t_0}{24} + \frac{t_1}{12} + \frac{t_0}{6} + \frac{\sigma_0}{4} + \frac{\sigma_1}{2} \right)$$

où E est la tension correspondant au niveau logique 1 des bits d'entrée.

Cette disposition permet un approvisionnement en éléments résistifs plus aisé que pour le montage d'addition pondérée équivalent, où la valeur des résistances d'entrée est inversement proportionnelle aux poids.

La valeur de R peut être choisie quelconque, la précision du convertisseur n'étant influencée que par la dispersion des éléments autour de la valeur commune et par les variations des niveaux logiques d'entrée.

Le calcul de l'erreur relative maximum sur la tension de sortie S_A du convertisseur donne :

$$\frac{\Delta S_A}{S_A} = \frac{\Delta E}{E} + \frac{\Delta R}{R}$$

Le nombre de niveaux à distinguer correspondant au nombre de quarts de ton qui composent les 4 octaves de la plage de mesure est de 96. Le pas de conversion vaut donc environ 1 % de la tension de sortie maximum. L'erreur commise sur S_A doit être suffisamment faible pour éviter la confusion entre deux valeurs voisines de la mesure : nous l'avons choisie égale à 0,2 % (ou 1/5 de pas). En prenant, pour le réseau d'addition, des résistances à 0,1 %, les niveaux logiques d'entrée doivent également être identiques à 0,1 % près. Cette précision n'étant pas atteinte par les sorties directes du registre tampon, il est nécessaire d'utiliser des portes analogiques qui commutent une tension de référence stabilisée.

Le compromis entre le nombre de niveaux à représenter et l'excursion maximale de la tension de sortie de l'amplificateur opérationnel nous a fait fixer le pas de conversion à 100 mV pour un quart de ton. Le domaine de variation de S_A , dans toute la plage de mesure est donc 0, -9,6 V.

Le choix d'une tension de référence de 5 V pour les entrées, facilement réalisable à l'aide d'un régulateur intégré, impose un rapport R'/R d'environ 2 ; l'impédance de chaque entrée étant supérieure à R ($1,5 R$ pour les q_i et t_i et R pour les σ_i) nous avons pris $R = 100 \text{ k}\Omega$ pour charger aux minimum le générateur de référence et $R' = 180 \text{ k}\Omega$ en série avec une résistance ajustable de $50 \text{ k}\Omega$.

3.4.4.5. *Dispositif d'initialisation*

La séquence de mesure est déclenchée par une impulsion du signal d'entrée. Il faut, à cet instant, réaliser l'initialisation générale de l'appareil pour le rendre prêt à la mesure. Trois impulsions, prélevées successivement sur l'horloge de période $t/100$ sont nécessaires :

- la 1ère (N_1) remet les compteurs et bistables à zéro, à l'exception du compteur de sortie ; la mesure peut alors commencer immédiatement ;
- la 2ème (N_2) assure le transfert du contenu (k) du compteur d'intervalles élémentaires dans le registre tampon ;
- la 3ème (N_3) remet le compteur de sortie à zéro, qui ne commencera à s'incrémenter qu'au bout du temps T_0 .

Les absences de mélodie, dues à des silences ou à des phonèmes non voisés, doivent être détectées sans ambiguïté. Cette fonction est réalisée par un circuit qui fournit une impulsion lorsque le contenu du compteur d'intervalles élémentaires dépasse sa capacité ; cette impulsion sert à mettre la valeur de la sortie à zéro qui sera maintenue ainsi jusqu'à la mesure suivante.

3.4.4.6. *Conclusion*

Le dispositif original de mesure logarithmique de fréquence instantanée que nous avons réalisé et implanté sur 5 petites

cartes de circuit imprimé logeables dans un boîtier peu encombrant (105 x 160 x 210 mm) ; il utilise, au total, 30 boîtiers de composants intégrés de type TTL. Son alimentation interne, stabilisée à partir du secteur alternatif, fournit une puissance de 3,5 W ; cette consommation pourrait être considérablement réduite par l'utilisation de circuits C MOS.

Le coût modéré de l'appareil, allié à son faible encombrement et à sa mise en oeuvre aisée, en fait un élément qui peut s'intégrer facilement à une chaîne d'analyse.

La connexion de la sortie "alternative" de l'électroglottographe à l'entrée du système de mesure de fréquence instantanée réalise l'analyseur de mélodie.

Aux chapitres suivants, nous précisons, à partir de nos expériences, les domaines d'application de cet ensemble qui sont :

- analyse de la parole (mélodie, intonation, micromélodie) ;
- synthèse de la parole (source de mélodie pour un vocoder) ;
- acoustique musicale ;
- étude des langues à tons.

C H A P I T R E 4

EXPERIMENTATION ET APPLICATIONS DE LA CHAÎNE

D'ANALYSE DE MÉLODIE

4.1. LA COURBE DE MÉLODIE

4.1.1. L'analyseur de mélodie

L'étude que nous avons retracée dans les chapitres précédents, a conduit à la réalisation d'un analyseur de mélodie d'un type nouveau par rapport à ceux qui ont été mis au point jusqu'à présent.

Nous rappelons ici son principe général qui se divise en deux parties :

- détection du mouvement des cordes vocales : elle fournit un signal électrique, image des impulsions laryngées, dont la composante fondamentale instantanée est prépondérante ; cette fonction est réalisée par l'électroglottographe ;
- mesure de la fréquence instantanée de ce signal : la transformation du glottogramme en une suite d'impulsions intervenant au début de l'ouverture de la glotte, permet de déterminer les durées des cycles glottiques successifs ; la valeur de la mesure, renouvelée à chaque période, est exprimée dans une échelle logarithmique, et est disponible sous forme numérique et analogique.

La mise en oeuvre du dispositif d'analyse est très simple, puisqu'elle ne nécessite que la pose des électrodes, préalablement enduites de pâte conductrice, sur le cou du locuteur, et le réglage, rapide, de l'électroglottographe ; les connections à réaliser avant l'expérience sont :

- branchement du collier porte électrodes ;
- liaison de la sortie "alternative" de l'électroglottographe à l'entrée de l'appareil de mesure de fréquence instantanée ;
- branchement de la sortie de mesure sur l'enregistreur.

L'étalonnage en unités logarithmiques de fréquence du support d'enregistrement peut être fait en connectant à l'entrée du système de mesure un générateur de signaux carrés de niveaux 0 et +5 V (niveaux logiques T.T.L) et de fréquence ajustable ; la valeur de la mesure est contrôlée sur les digits d'affichage.

4.1.2. Présentation des mesures

La sortie numérique de l'analyseur de mélodie, matérialisée par les 7 bits du mot {O.T.Q.} est essentiellement destinée au chargement des mesures successives dans une mémoire digitale telle que celle d'un ordinateur, en vue d'applications spéciales comme, par exemple, la commande de la source vocale d'un vocoder (cf. § 4.3).

Mais à chaque fois que l'on s'intéressera à la visualisation du schéma mélodique de la parole, on utilisera la sortie analogique qui fournit une tension proportionnelle au logarithme du rapport de la fréquence instantanée mesurée (F) à la fréquence de référence (F_0). Cette tension, appliquée à l'entrée d'un enregistreur graphique rapide (bande passante au moins égale à 1000 Hz), permet d'obtenir directement la courbe mélodique de la phrase.

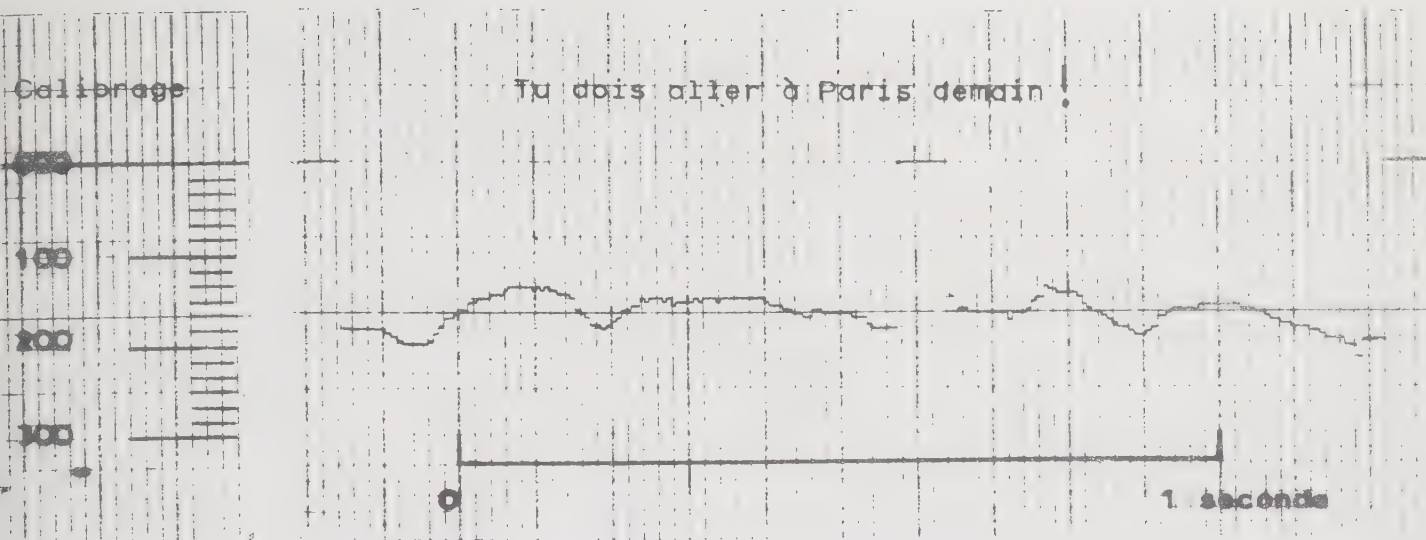


Figure 51 - EXEMPLE DE COURBE MELODIQUE .

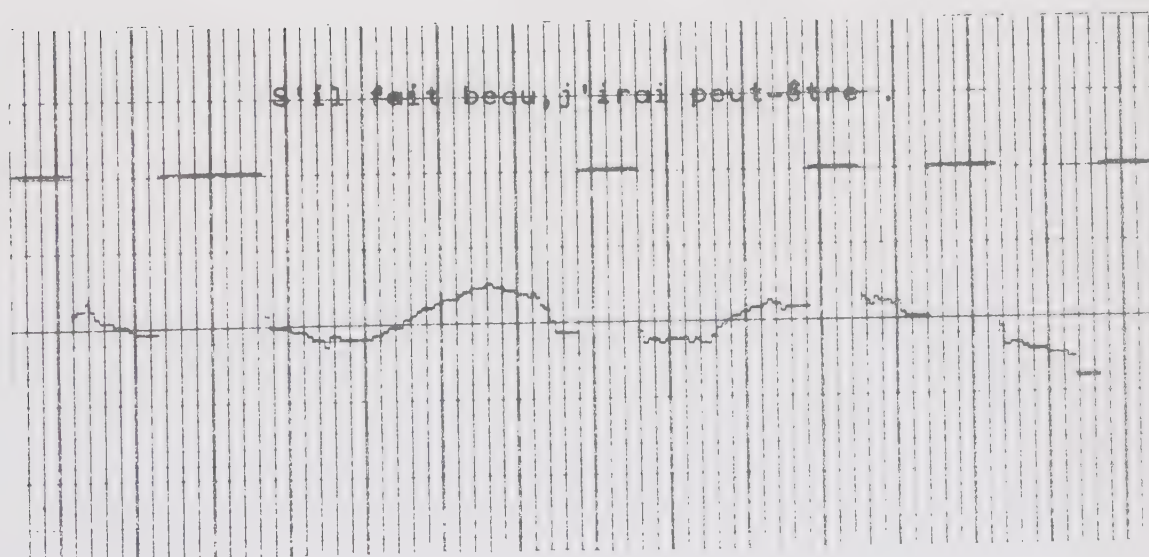


Figure 52 - EXEMPLE DE COURBE MELODIQUE .

Les figures 51 et 52 montrent le type de courbes que l'on obtient sur un enregistreur à jet d'encre (Oscillomink) à l'aide de notre analyseur. La ligne supérieure, interrompue pendant la durée de vibration des cordes vocales représente l'indication "000" correspondant à la fréquence de référence ou à l'absence de mélodie ; on y remarque, en particulier, la présence des consonnes non voisées ($|f|$ et $|p|$). L'étalonnage en octaves et tons permet d'apprécier directement les différences de hauteur.

4.2. ANALYSE DES VARIATIONS DE HAUTEUR DE LA PAROLE

4.2.1. Introduction

Les variations de hauteur dans la voix parlée, que notre analyseur permet de visualiser, résultent de la superposition de plusieurs des composantes mélodiques possibles suivantes :

- les faits d'origine physiologique, influant sur le nature de la voix :

* effets mélodiques propres au locuteur (ou à certaines catégories de locuteurs : "accent" local) ;

* micromélodie ;

- les faits prosodiques, chargés d'information linguistique :

* accent ;

* intonation ;

* tons.

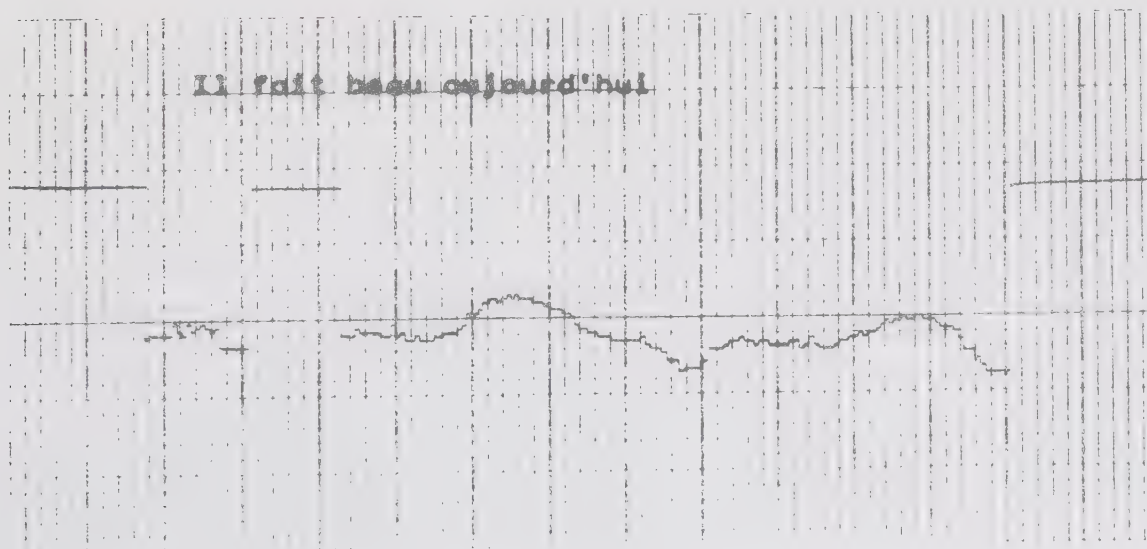
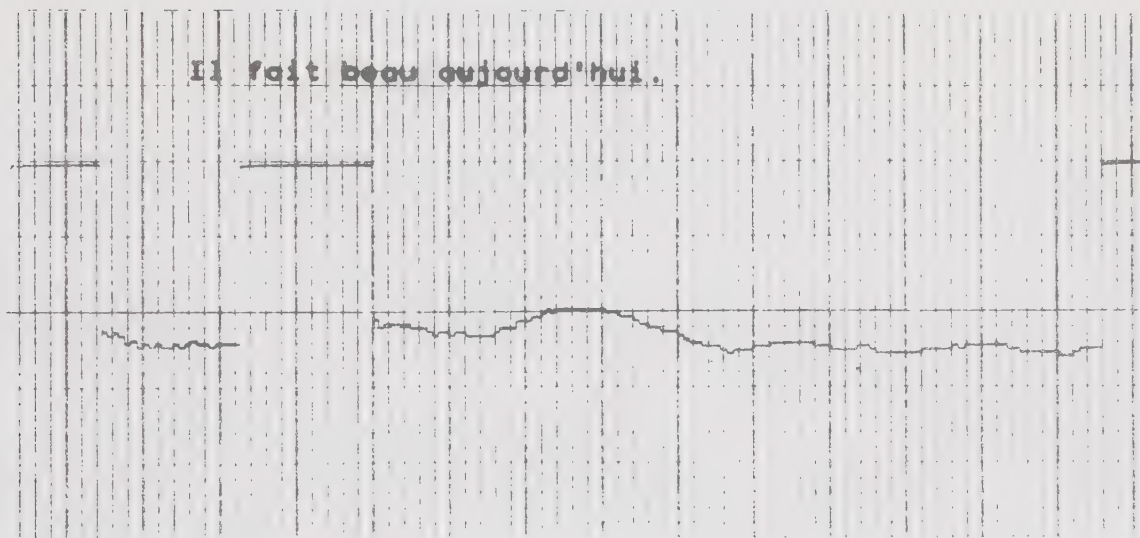
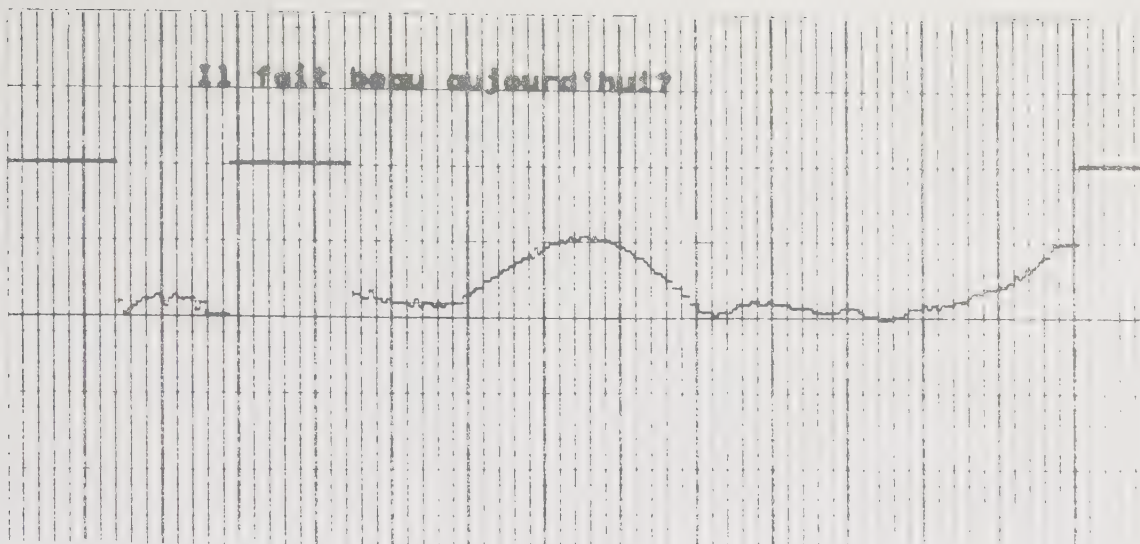


Figure 53 - ETUDE D'INTONATION .

La multiplicité des informations présentes simultanément dans la mélodie, fait de l'analyse séparée de chaque composante une opération délicate du point de vue de l'interprétation de la courbe mélodique. Elle demande une préparation sérieuse du corpus et un choix judicieux du locuteur, en fonction de la variable à mettre en évidence.

Dans les paragraphes suivants nous suggérons, à titre d'exemple, l'utilisation de notre dispositif pour la détermination des variations de hauteur dues à l'intonation et à la micromélodie, puis, au dernier chapitre, nous développerons plus spécialement le cas des langues à tons.

4.2.2. Détermination de schémas intonatifs simples

L'intonation concerne l'agencement mélodique global d'un segment de parole et permet d'en adapter le sens.

Ainsi, la différence entre l'interrogation, l'affirmation et l'affirmation brutale apparaît sur les enregistrements des courbes mélodiques (figure 53) concernant la phrase :

"Il fait beau aujourd'hui".

Le schéma mélodique général diffère dans les trois cas, principalement à la fin des phrases où la forme de la variation de hauteur constitue le message intonatif :

- interrogation : hauteur montante ;
- affirmation : hauteur constante ;
- affirmation brutale : hauteur descendante.

Il est à remarquer que ces trois types d'intonation ne constituent pas des niveaux significatifs, et qu'il existe, entre eux, une quantité de nuances qui font

apparaître graduellement le doute depuis l'affirmation brutale jusqu'à l'interrogation.

4.2.3. Micromélogie

Nous avons défini la micromélogie (§ 1.3.2) comme étant une légère altération de la ligne générale des variations de hauteur de la parole dans les phonèmes vocaliques.

Ce phénomène incontrôlable, d'origine essentiellement physiologique, apparaît au niveau des liaisons voyelle-consonne, et se manifeste aussi bien sur les voyelles que sur les consonnes voisées ; il est indépendant de la mélodie du message.

L'importance de la perturbation de hauteur due à la micromélogie varie avec les locuteurs, et peut atteindre 20 % de la valeur du phonème vocalique considéré [54] . La précision de mesure de notre analyseur (± 2 %) nous interdit l'analyse fine de la micromélogie. Une deuxième version de l'appareil de mesure de fréquence instantanée [2], actuellement en construction au laboratoire, devrait permettre d'aborder quantitativement ce type d'analyse, par accroissement de la précision. La quantification est portée à 1/8 de ton et l'erreur due à l'approximation des dates T_p inscrites dans la table séquentielle à 0,25 %. L'erreur relative sur la période mesurée est alors de 1 % au maximum.

La durée du phénomène de micromélogie est très brève (quelques périodes laryngées). Son observation est délicate si l'on ne dispose pas d'un système l'isolant efficacement dans le temps de façon à pouvoir le dilater sur l'enregistrement. Ce problème peut être résolu par l'emploi d'un segmentateur qui permet de sélectionner à volonté des portions de signal à l'intérieur d'une phrase.

Nous avons réalisé un appareil de ce type, dont la description est donnée à l'annexe I.

4.3. SYNTHESE DE LA MELODIE DE LA PAROLE

4.3.1. L'analyse-synthèse de la parole par vocoder

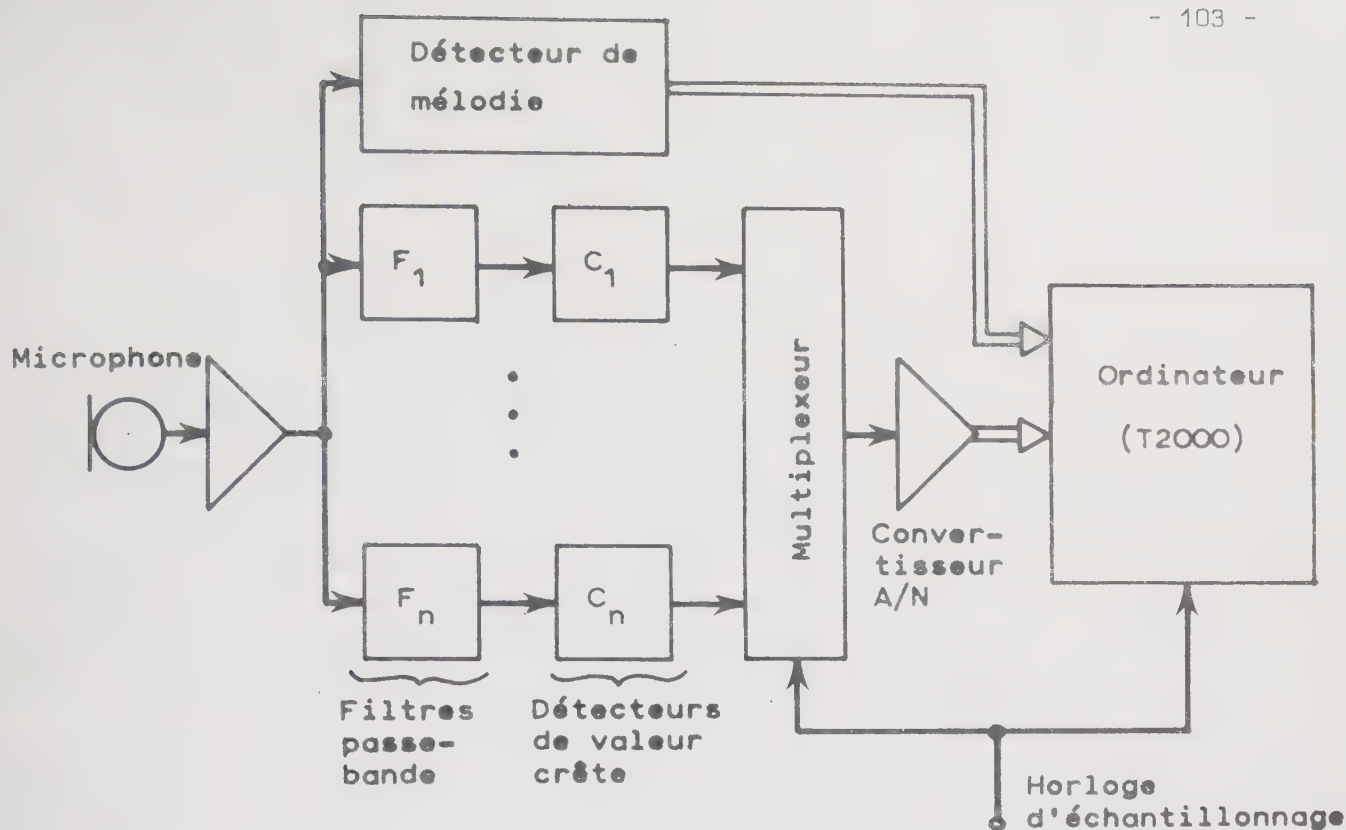
Le vocoder est un appareil qui peut être classé parmi les synthétiseurs de parole. Sa commande est réalisée par les paramètres spectraux d'un signal préalablement analysé.

L'apparition du vocoder est liée à la nécessité de réduire la quantité d'information parlée transmise sur les lignes de télécommunication : il permet, en effet, de n'utiliser, à la synthèse, qu'une partie des caractéristiques acoustiques redondantes de la parole, tout en conservant l'intelligibilité.

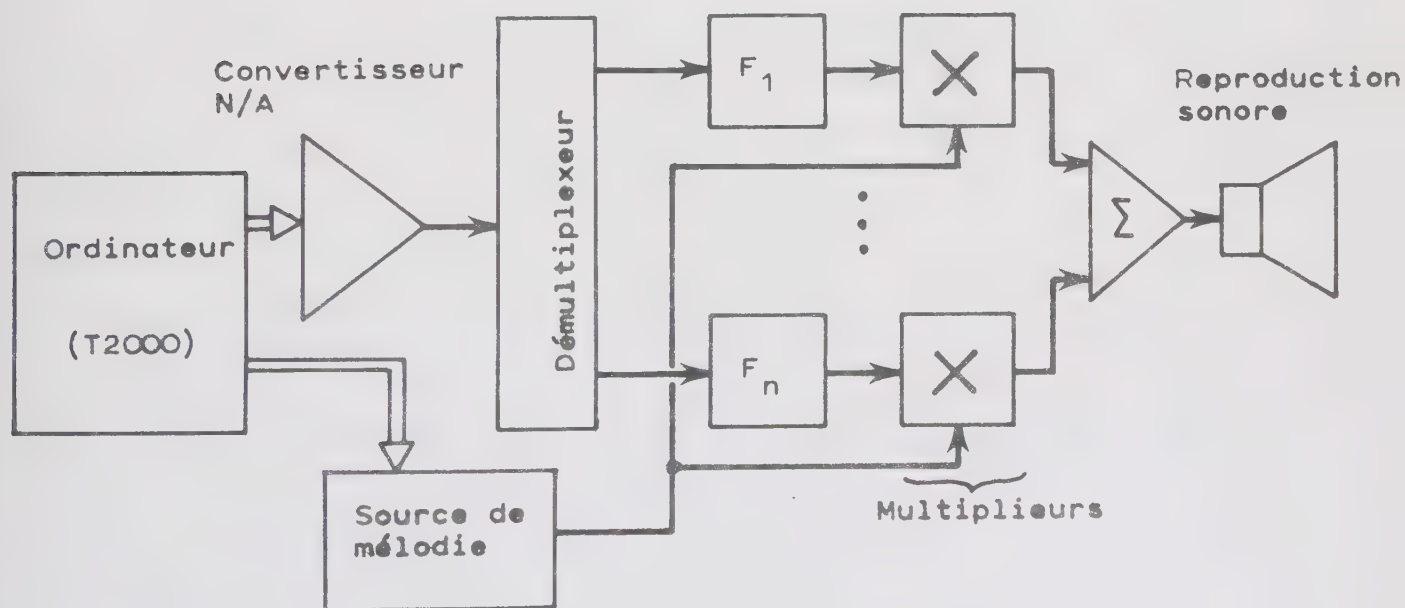
Dans le vocoder actuellement en construction au laboratoire d'Automatique de Besançon, la ligne de transmission est remplacée par la mémoire d'un ordinateur, ce qui permet, après la phase d'analyse, de réaliser autant de synthèses qu'on le désire, avec la possibilité de modifier les paramètres enregistrés. Nous obtenons de la sorte un synthétiseur de parole qui, sans prétendre aux performances des synthétiseurs à formants, permet de réaliser un grand nombre d'expériences sur la parole.

Le vocoder se compose essentiellement de deux parties (figure 54) :

- à l'analyse, un ensemble de filtres passe-bande à accords décalés les uns par rapport aux autres, et disposés en parallèle, permet d'extraire l'amplitude instantanée des composantes harmoniques du signal de parole et la fréquence mélodique ; l'information enregistrée est le spectre évolutif discrétisé du message dont la durée, limitée par un compromis entre la taille de



a) Configuration "analyseur"



b) Configuration "synthétiseur"

Figure 54 - SCHEMA SYNOPTIQUE DU VOCODER .

la mémoire et la fréquence d'échantillonnage, peut atteindre 2 s ;

- à la synthèse, un banc de filtres analogues à celui de l'analyse, est excité par un signal riche en harmoniques, et dont la fréquence instantanée est commandée par le paramètre de mélodie ; l'amplitude des composantes sur chaque voie est déterminée à tout instant par les valeurs enregistrées ; la sommation des sorties des filtres permet de reconstituer un signal dont le spectre est voisin de celui de la parole originale, mais moins complet.

Le point le plus délicat dans la réalisation du vocoder est la détection et la restitution de la mélodie.

Nous avons donc pensé à adapter notre analyseur à ce problème.

4.3.2. Détection et reproduction de la mélodie

Dans la phase d'analyse, notre dispositif permet, par sa sortie numérique de la mesure, d'enregistrer la valeur de la fréquence instantanée des productions laryngées, simultanément avec les paramètres d'amplitude des composantes spectrales du signal de parole.

A la synthèse, il faut disposer d'une source de signal riche en harmoniques à fréquence commandée. Celle-ci est obtenue par transformation simple du dispositif de mesure de fréquence instantanée en générateur (figure 58) ; il suffit de comparer le contenu du compteur de sortie aux valeurs enregistrées lors de l'analyse et d'envoyer les impulsions de coïncidence sur l'entrée du circuit d'initialisation ; la fréquence de ces impulsions est fixée par la valeur de la commande appliquée au comparateur et varie comme elle. On obtient un signal carré au lieu d'une suite d'impulsions en faisant agir celles-ci sur l'entrée de basculement d'un bistable JK

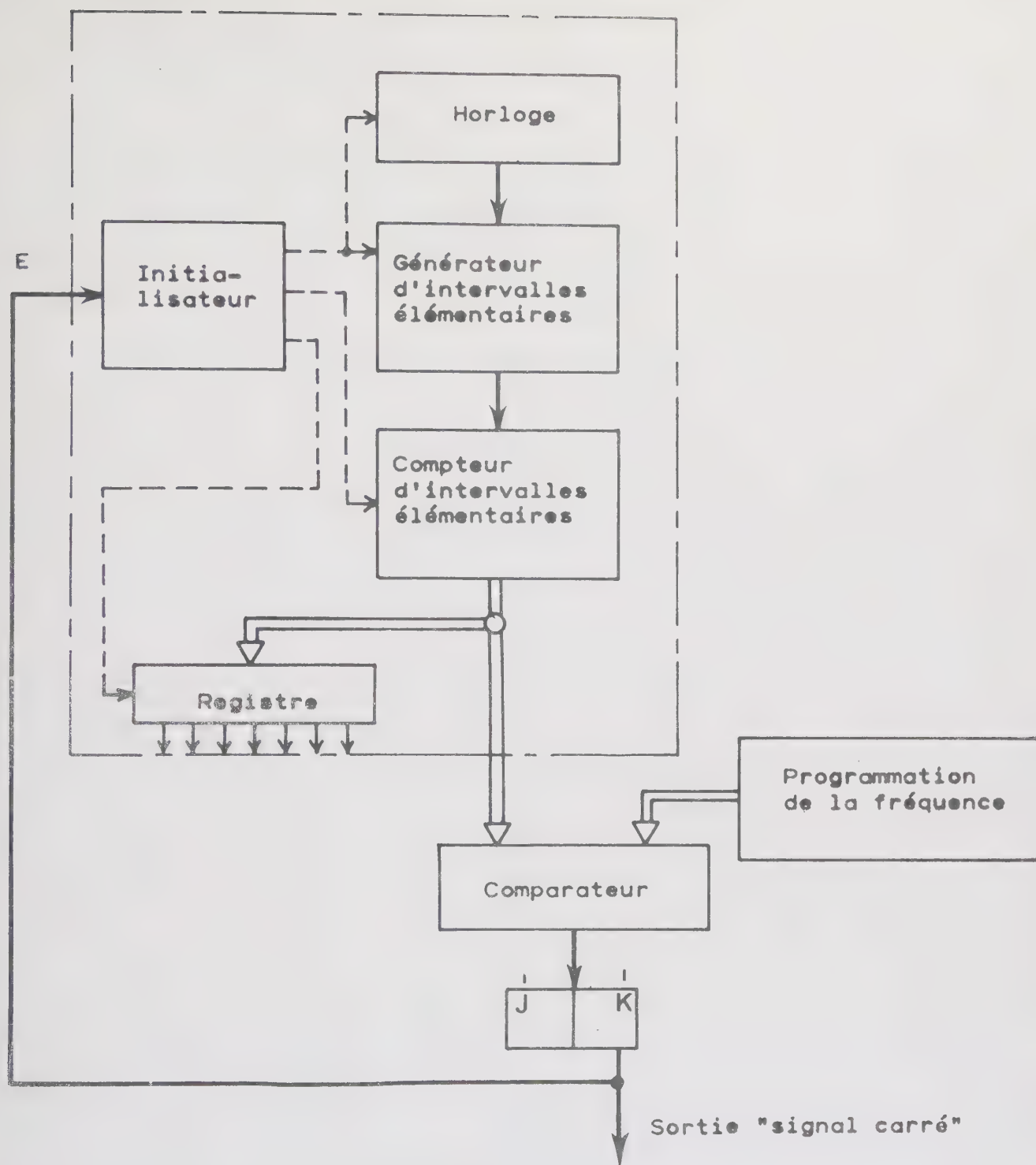


Figure 55 - SCHEMA DE PRINCIPE DU GENERATEUR DE MELODIE PROGRAMMABLE .

après avoir multiplié la fréquence d'horloge par 2.

Ce type de vocoder, fonctionnant avec une mélodie précise (incluant éventuellement la micromélodie), doit permettre d'obtenir une parole synthétique de bonne qualité et aisément modifiable par action sur les paramètres enregistrés.

4.3.3. Autres applications du générateur de mélodie

4.3.3.1. *Générateur des notes de la gamme musicale*

Le générateur de mélodie [2] est actuellement en cours de fabrication au laboratoire ; il est quantifié en huitièmes de ton.

La production du signal de sortie est parfaitement synchronisée avec l'horloge. La précision du générateur est donc fixée uniquement par l'approximation des dates T_p inscrites dans la table séquentielle : l'erreur maximum relevée est de 0,373 %, soit environ $1/32$ de ton.

Le compteur d'intervalles élémentaires peut être construit de façon à faire apparaître les douze notes de la gamme musicale dont les fréquences, programmables manuellement (par des disques codeurs), s'échelonnent tous les demi-ton ; la précision d' $1/32$ de ton est conservée grâce au fonctionnement synchrone du dispositif avec l'horloge. On constitue ainsi un générateur de notes de référence pouvant servir, par exemple, au contrôle de l'accord des instruments de musique.

4.3.3.2. *Générateur de signaux à variation de fréquence logarithmique*

Le générateur de mélodie est aussi programmable dynamiquement, comme c'est le cas pour la source du vocoder. Associé à un convertisseur assurant la transformation de la suite d'impulsions ou de l'onde carrée de sortie en un signal de forme donnée, et par exemple sinusoïdale, il peut servir à exciter des systèmes vibrant en basse fréquence en vue d'en déterminer la courbe de réponse ; le balayage logarithmique programmé à pas fin ($1/8$ de ton) permet d'analyser avec précision et en une seule fois une large bande de fréquence.

L'obtention d'une onde de forme donnée, peut notamment se faire de deux façons :

- filtrage passe-bande à fréquence centrale asservie du signal carré : on peut obtenir ainsi des signaux sinusoïdaux de même fréquence (jusqu'à 16000 Hz et dans une plage de 10 octaves en dessous, dans notre cas) ;
- génération numérique de signaux par comptage des impulsions : ce procédé permet d'obtenir de nombreuses formes de signaux, mais réduit considérablement les fréquences utilisables (la définition de la période par 100 échantillons, par exemple, divise les fréquences disponibles par 100 ; nous pourrions obtenir, dans ce cas, un maximum de 320 Hz qui est suffisant dans de nombreuses applications de la mécanique des matériaux et des structures).

4.4. CONCLUSION

Le dispositif d'analyse de mélodie, dont ce mémoire constitue l'étude, est particulièrement adapté à la détermination en temps réel des variations de hauteur de la voix.

Le prototype que nous avons construit nous a permis de réaliser de nombreuses expériences, dans lesquelles nous avons testé sa bonne fiabilité et sa simplicité d'emploi.

Il permet d'analyser aisément l'intonation, les variations de hauteur dues aux faits accentuels ou aux traits propres au locuteur, et, par amélioration de la précision de mesure, la micromélodie. Nous pouvons même envisager, à l'inverse, la synthèse de la mélodie pour réaliser les sources vocales des synthétiseurs de parole.

Parmi les applications originales du générateur de mélodie, nous avons signalé, dans le domaine de l'acoustique musicale, la production de notes de référence, et, pour l'instrumentation scientifique, la génération de signaux à variation logarithmique programmable de fréquence.

Dans les applications linguistiques de notre chaîne d'analyse de mélodie, il reste notamment l'étude des variations de hauteur dans les voyelles des langues à tons que nous avons réservée pour le dernier chapitre.

C H A P I T R E 5

L'ANALYSE DES LANGUES A TONS

5.1. LES LANGUES A TONS

L'usage linguistique des variations de hauteur mélodiques dans des langues telles que le français ou l'anglais, est essentiellement limité à la production de l'intonation et des faits accentuels.

Mais dans d'autres langues, la mélodie est utilisée à des fins distinctives, à l'aide d'unités discrètes : ce sont les langues à tons qui se rencontrent principalement en Asie de l'Est et du Sud-Est, en Afrique (au Sud du Sahara et en Ethiopie) en Amérique du Nord (dans les peuplades indiennes des USA et du Mexique).

Les tons sont classés parmi les faits prosodiques, car un ton donné peut affecter différents phonèmes et même, parfois, un segment composé de plusieurs phonèmes, comme c'est le cas, par exemple, de quelques mots danois, suédois et norvégien.

Un mot d'une langue à tons est donc parfaitement défini lorsque l'on a identifié à la fois ses tons et ses phonèmes.

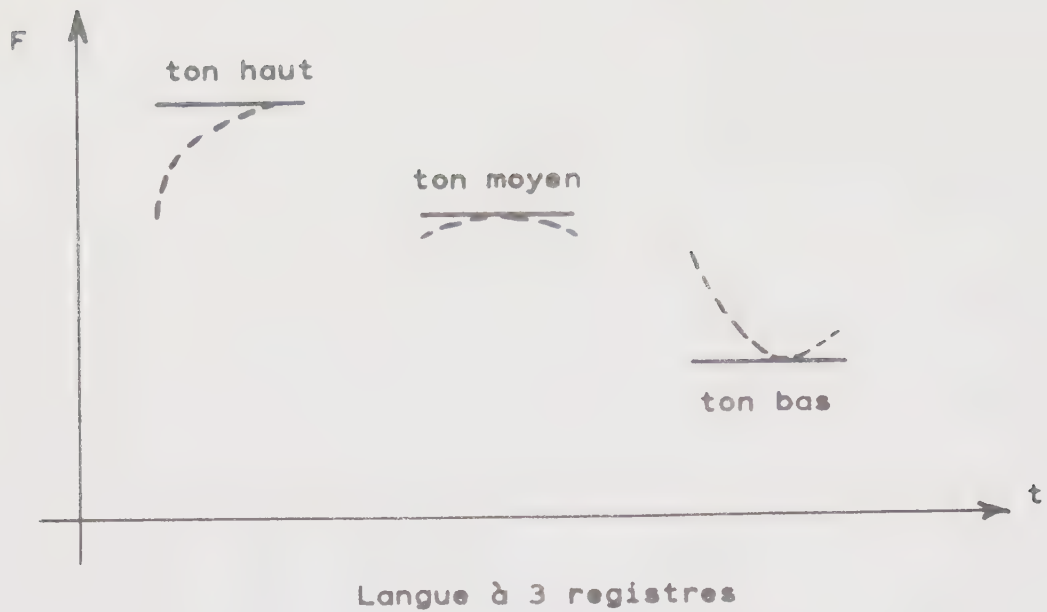


Figure 56 - TONS PONCTUELS .

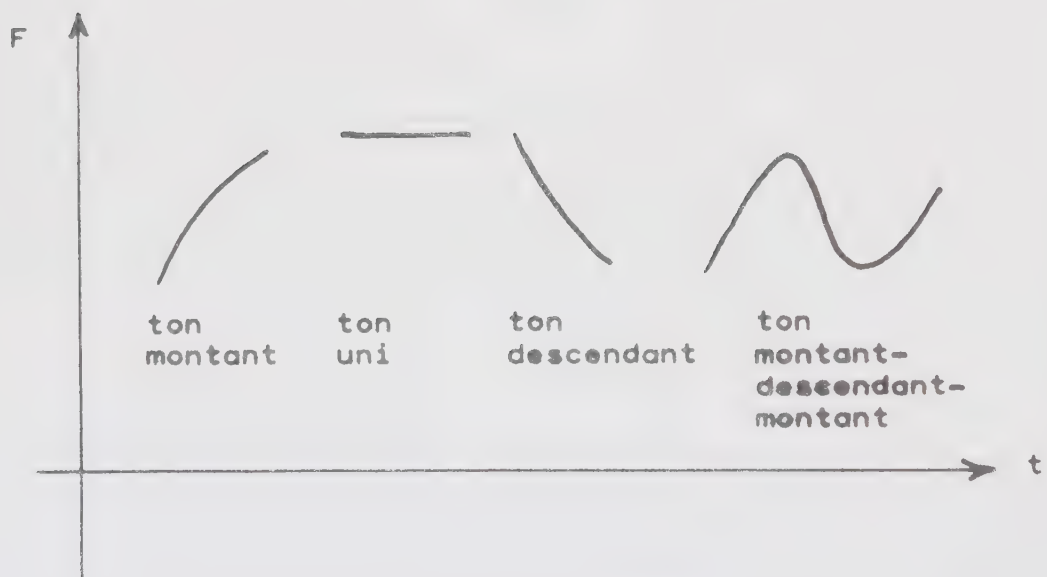


Figure 57 - TONS MELODIQUES .

Un ton est toujours caractérisé par un mouvement mélodique de longueur variable ; mais ce mouvement n'est pas obligatoirement utilisé pour différencier un ton d'un autre dans une langue donnée.

On distingue principalement deux types de systèmes tonaux :

- les langues à tons ponctuels dans lesquelles n'est significatif qu'un point de la mélodie de chaque ton, caractérisant son niveau ou registre ; un ton ponctuel ne peut être reconnu que par comparaison avec les syllabes voisines ; le nombre de registres utilisés (figure 56) varie avec les langues et peut être de 2 (ton haut et ton bas), 3 (ton haut, ton moyen, ton bas) ou plus rarement 4 ;
- les langues à tons mélodiques où le trait distinctif est la direction de la ligne mélodique des tons ; on peut y trouver des tons à direction unique (montants, descendants, uniformes) et des tons à un ou plusieurs changements de direction (figure 57).

La présence des tons dans une langue n'implique pas une absence d'intonation. Celle-ci se superpose à la mélodie des tons ; il peut arriver, par exemple, que dans deux mots différents d'une même phrase le ton haut de l'un soit plus bas que le ton bas de l'autre, et inversement ; mais seul compte le respect du schéma tonal qui peut être déplacé globalement vers le haut ou vers le bas, selon l'intonation, sans changer le sens du mot.

L'étude linguistique des langues à tons nécessite donc l'utilisation intensive d'un analyseur de mélodie. Notre dispositif, fiable, simple d'emploi et fournissant directement une indication logarithmique (en octaves, tons et quarts de ton) des différences de hauteur de la voix est, dans ce cas, un outil particulièrement intéressant.

5.2. ETUDE DES TONS DU YORUBA

Dans le cadre d'un travail effectué par M. Davio au laboratoire de phonétique de Besançon, nous avons pu disposer d'un locuteur dont la langue maternelle est le Yoruba.

Le Yoruba est une langue à tons ponctuels, parlée dans l'Ouest du Nigéria et dans l'Est du Dahomey ; elle est caractérisée par un système à 3 registres.

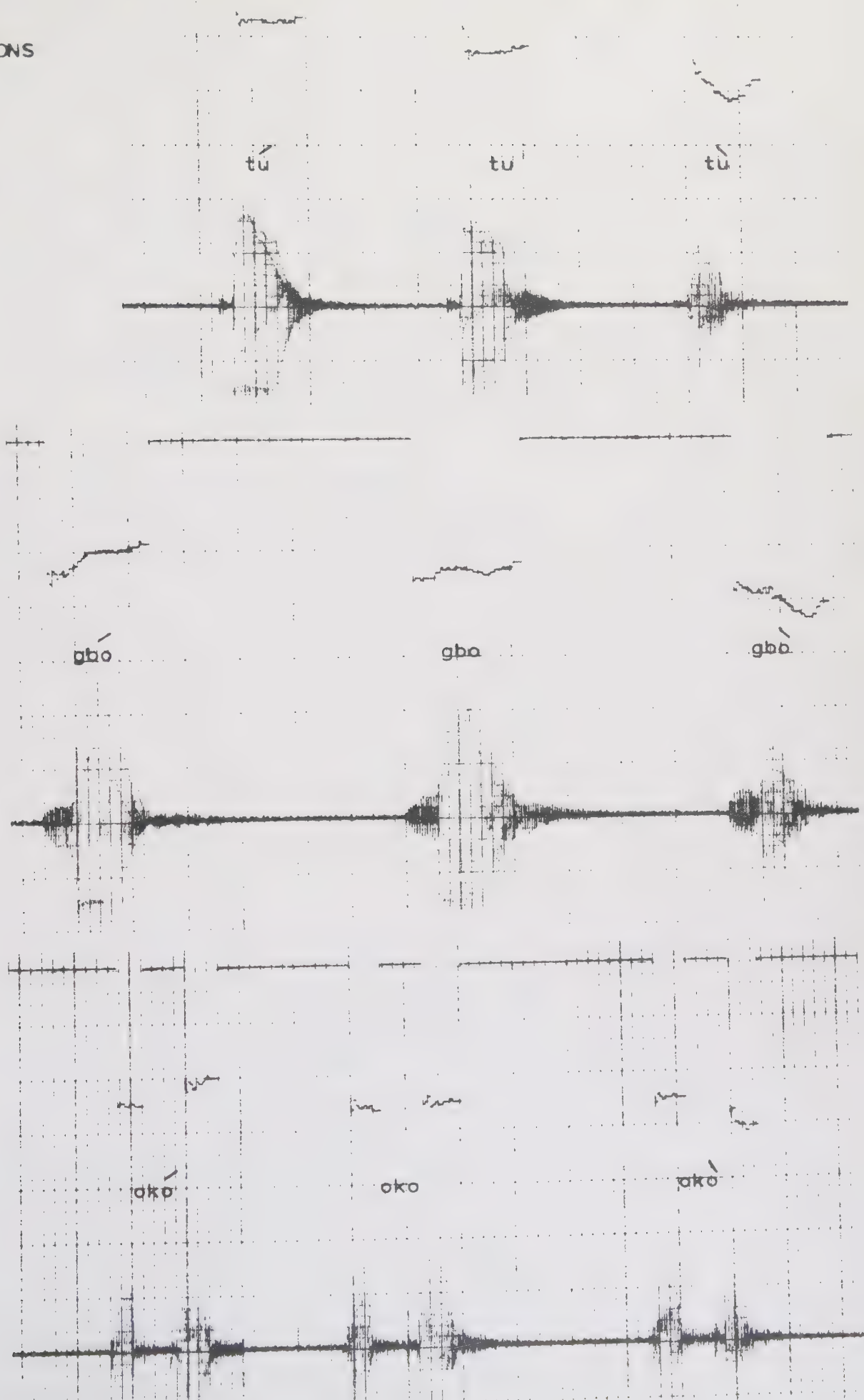
Le problème consistait à analyser la mélodie et la durée des tons de la langue en fonction de la configuration des mots. Le corpus utilisé faisait apparaître :

- des syllabes à un ton formées de consonnes et de voyelles diverses, avec variation du type de ton (ton haut, moyen ou bas) ;
- des segments de deux syllabes composés de deux tons séparés par une consonne ; on y faisait varier l'opposition de deux types de tons ;
- des phrases faisant apparaître l'opposition de deux tons successifs au sein d'un message complet.

La partie expérimentale de l'analyse a été effectuée en deux phases :

- enregistrement, en une fois, des glottogrammes et des phonogrammes sur les deux pistes d'un enregistreur magnétique NAGRA IV S, pour la totalité du corpus ;
- tracé sur un enregistreur graphique à jet d'encre (OSCILLOMINK) des courbes mélodiques obtenues en reliant directement la sortie de l'enregistreur à l'entrée de l'appareil de mesure de fréquence instantanée.

Figure 58 -
EXEMPLE DE TONS
EN YORUBA



Le dépouillement en temps différé permet de ne pas immobiliser le locuteur trop longtemps dans le cas d'un long corpus.

Les figures 58a à 58t reproduisent quelques enregistrements de chaque partie du corpus ; ils montrent, en particulier, l'aptitude du système à restituer les mouvements mélodiques à l'intérieur des tons.

5.3. DISPOSITIF D'ENTRAINEMENT A LA PRODUCTION ET A LA PERCEPTION DES VARIATIONS TONALES DE LA VOIX

L'analyseur de mélodie que nous avons mis au point peut se transformer simplement en un système de visualisation en temps réel de la courbe de mélodie de la voix d'un locuteur.

La figure 59 en montre le principe qui consiste à relier la sortie de l'analyseur à l'entrée de l'amplificateur vertical d'un oscilloscope, dont le déclenchement du balayage est effectué de l'extérieur par le signal disponible à la sortie du circuit de mise en forme de l'appareil de mesure ; à chaque fois que le spot est revenu à sa position initiale, la mise en mouvement des cordes vocales provoque immédiatement le départ d'un nouveau balayage (calibré sur une faible vitesse).

L'utilisation d'un oscilloscope dont l'écran est divisé en deux parties dont l'une peut conserver une courbe modèle en mémoire, et l'autre fait apparaître la mélodie du locuteur en temps réel, autorise la comparaison visuelle immédiate et répètable de la production et de la référence.

Ce système permet de réaliser un apprentissage des schémas mélodiques plus efficace que par la chaîne naturelle : pour ses problèmes d'apprentissage d'activité vocale, l'homme dispose (figure 60, d'une chaîne de réaction constituée par une partie du cerveau qui commande l'appareil phonatoire à partir des données fournies par l'oreille ;

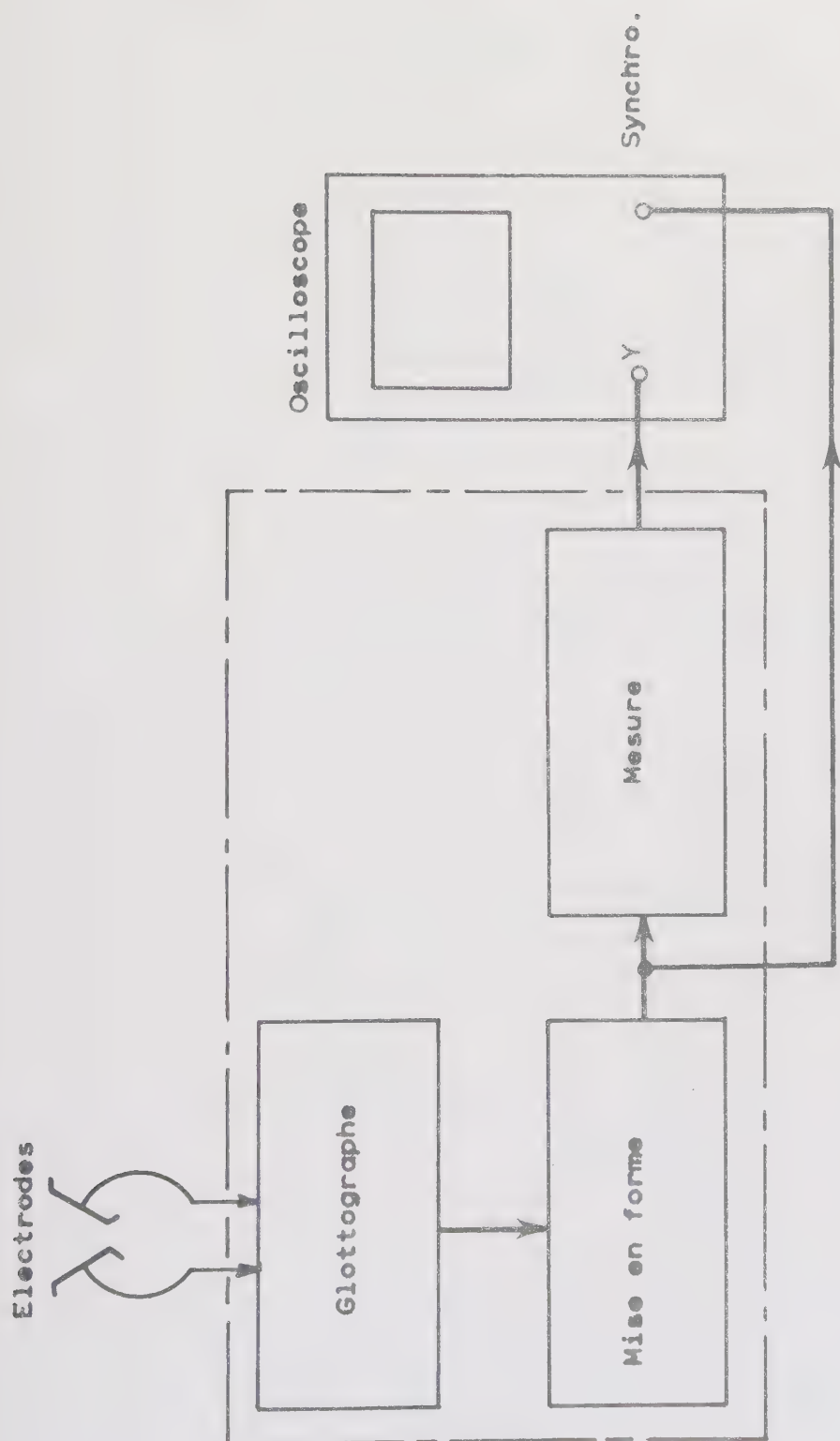
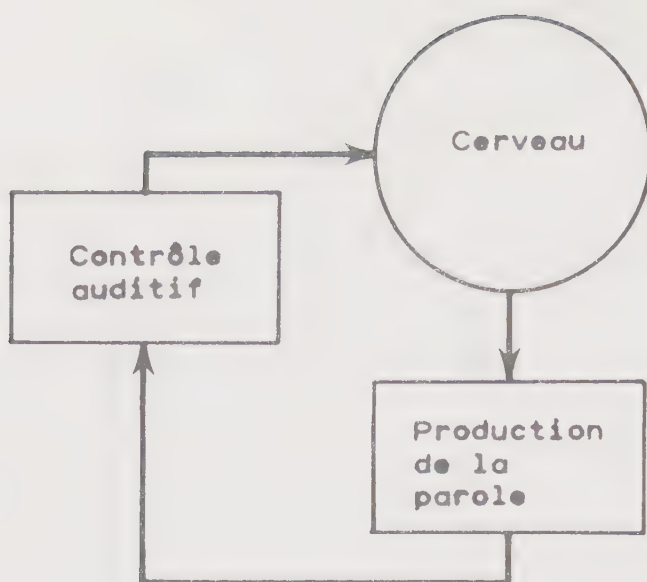
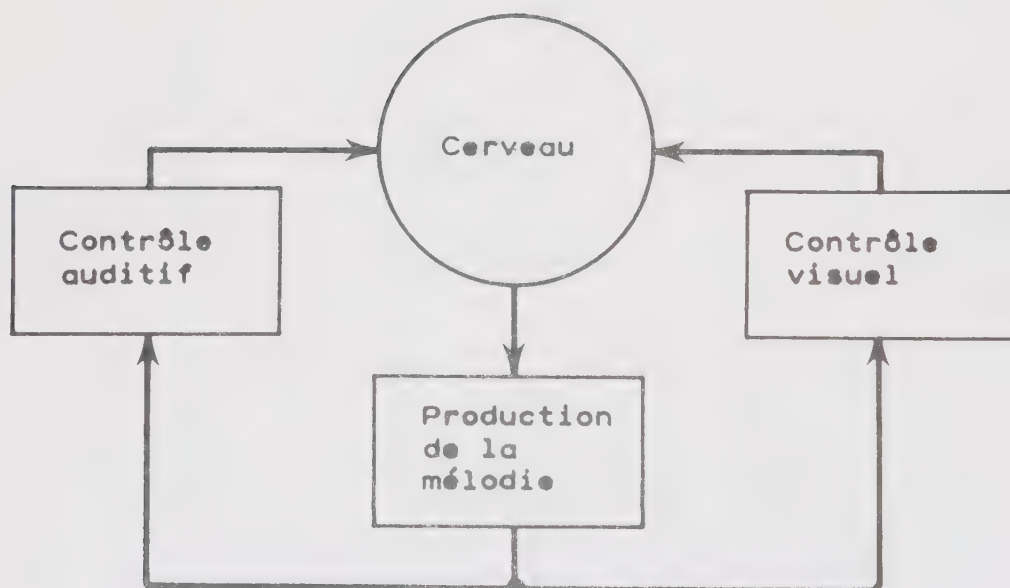


Figure 59 - DISPOSITIF DE VISUALISATION EN TEMPS REEL DE LA COURBE MELODIQUE .



a) Chaîne d'apprentissage naturelle



b) Adjonction d'une boucle visuelle pour l'apprentissage de la mélodie

Figure 60 - PRINCIPE DU PHENOMENE D'APPRENTISSAGE DE LA PAROLE .

mais, pour un locuteur non habitué qui s'exerce seul à imiter une mélodie fixée, le dispositif de production et le capteur auditif de contrôle sont très médiocres ; le rendement est alors très faible ; le système de visualisation en temps réel de la mélodie, associé à un gabarit de référence, réalise une deuxième chaîne de réaction (figure 67) passant par un contrôle visuel beaucoup plus précis et objectif ; nous pensons que le processus d'apprentissage de schémas mélodiques peut ainsi converger très rapidement comparativement à la chaîne naturelle.

Le même gabarit de mélodie peut être utilisé par tous les locuteurs, quelle que soit la hauteur moyenne de leur voix, puisqu'il est tracé dans une échelle logarithmique et représente des variations relatives de fréquence instantanée.

Nous disposons de la sorte d'un appareillage simple d'emploi pouvant servir à l'entraînement à la production et aussi à la perception (en remplaçant le locuteur apprenti par un locuteur de référence) des variations tonales et mélodiques de la parole. Les domaines d'application intéressés sont l'apprentissage des langues à tons (l'enseignement du chinois, par exemple, pourrait en bénéficier) et la familiarisation des étudiants avec les particularités intonatives et les faits accentuels de certaines langues.

A N N E X E I

REALISATION D'UN SEGMENTATEUR

1 - PRINCIPE

Le rôle du segmentateur au sein d'une chaîne d'analyse, est de ne transmettre le signal électrique appliqué à son entrée qu'entre deux instants (t_1 et t_2) préalablement choisis et repérés par rapport au signal.

Le principe du segmentateur consiste à multiplier la grandeur $s(t)$ représentant le message transmis, par une fonction $p(t)$ qui vaut 1 entre t_1 et t_2 , et zéro partout ailleurs. Cette fonction peut être réalisée au moyen d'un interrupteur dont l'ouverture et la fermeture sont commandées automatiquement aux instants prédéterminés.

Le segment obtenu est identique en amplitude au signal original, entre t_1 et t_2 , mais son spectre instantané est évidemment modifié par la présence de $p(t)$. Ce phénomène est gênant lorsque l'appareil est utilisé dans une chaîne d'analyse spectrale ; il conviendrait, dans ce cas, d'étudier la forme de l'ouverture et de la fermeture de l'interrupteur, donc des variations d'amplitude initiale et finale du segment, de telle sorte que ses perturbations spectrales soient définies. Cela revient à placer en série

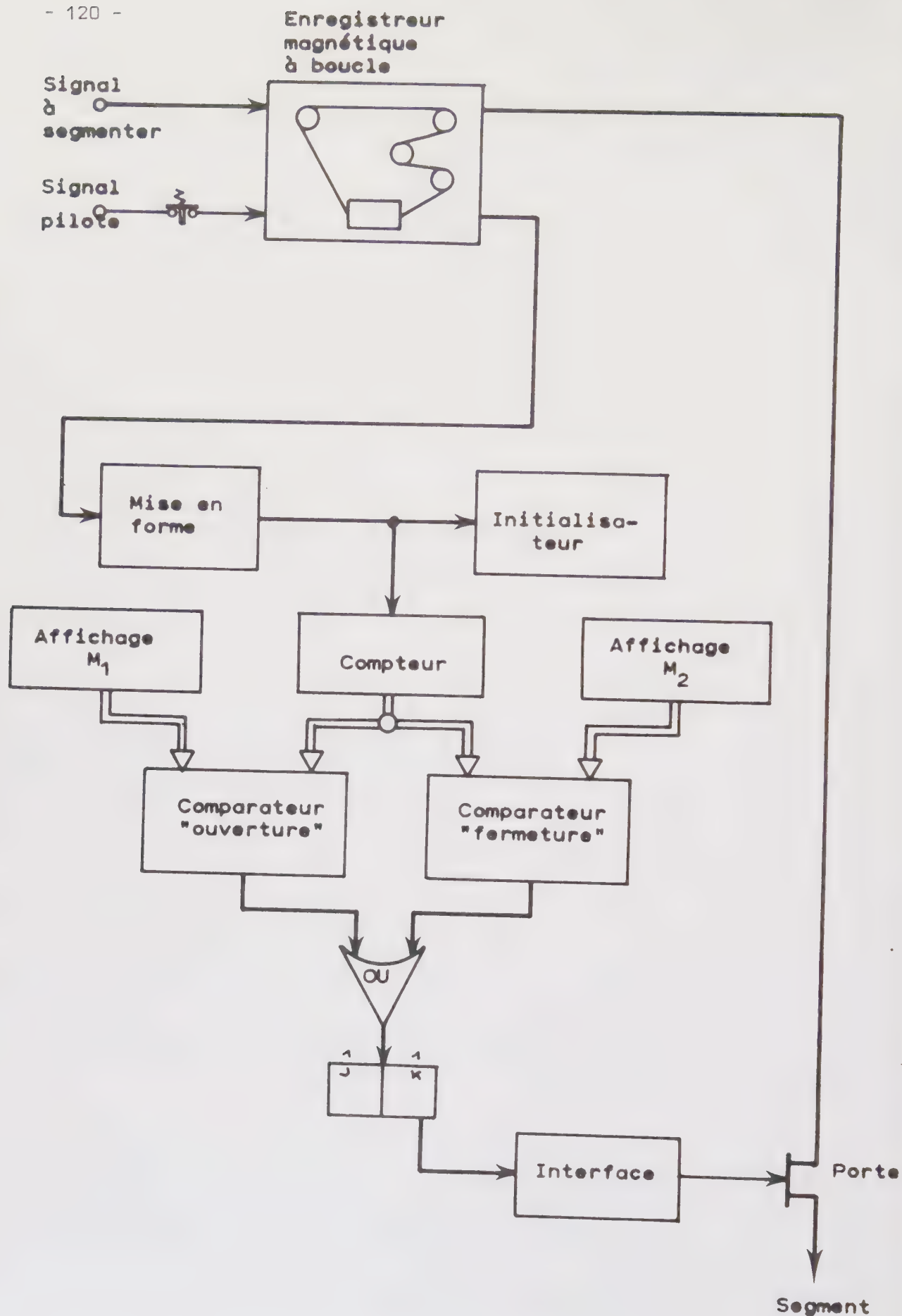


Figure 61 - SCHEMA SYNOPTIQUE DU SEGMENTATEUR .

avec l'interrupteur, des filtres ayant une courbe de réponse amplitude-fréquence adéquate.

Dans notre cas, l'analyse de mélodie de fractions de phrases ne souffre pas de la production d'harmoniques supplémentaires, puisque nous ne conservons que la composante d'ordre 1 du signal segmenté qui est ici le glottogramme.

Le fonctionnement du segmentateur, représenté par le schéma de la figure 61, se décompose en deux parties :

- le signal à segmenter est enregistré sur une piste d'une boucle de bande magnétique ; simultanément, un train d'impulsions à fréquence fixe est inscrit sur une deuxième piste, réalisant un échantillonnage temporel du signal ;
- à la reproduction, les impulsions provenant de l'enregistreur dont la boucle tourne "sans fin", sont dirigées vers la logique de commande de l'appareil qui les compte et compare leur nombre à deux valeurs M_1 et M_2 prédéterminées, fixant ainsi les instants d'ouverture et de fermeture d'une porte analogique qui sert d'interrupteur dans le circuit du signal ; le dispositif est initialisé automatiquement à chaque tour de la boucle, de façon à retrouver périodiquement les mêmes points de sélection.

2 - REALISATION

Nous nous sommes imposé, pour la construction du segmentateur, les caractéristiques suivantes :

- durée maximale du segment : 2,5 s ;
- pas de segmentation : 250 μ s ;

- admission d'un signal d'au moins 5 V d'amplitude ;
- bande passante : 0 à 20 kHz sans atténuation.

L'obtention d'une durée de 2,5 s par pas de 250 μ s nécessite une capacité de comptage de 10^4 (soit 4 décades).

La valeur du pas est obtenue par une horloge interne à 4000 Hz qui fournit les impulsions du signal pilote.

Le contenu du compteur est appliqué en permanence à deux comparateurs recevant respectivement la valeur des nombres d'impulsions M_1 et M_2 représentant les instants d'ouverture et de fermeture de la porte, codée au moyen de modules d'affichage manuels. Chaque comparateur est composé de 4 circuits intégrés en cascade réalisant directement la fonction. Les deux signaux de coïncidence (n_1 et n_2) sont envoyés sur l'entrée d'un bistable JK dont la sortie constitue la variable de commande de la porte analogique.

Le circuit d'initialisation est basé sur le fait que le signal pilote n'est pas enregistré sur la totalité de la longueur de la boucle. Il fournit une impulsion de remise à zéro du compteur et du bistable lorsque la durée entre deux fronts "descendants" du signal pilote est supérieure à 350 μ s.

La porte analogique est constituée d'un transistor à effet de champ canal N étudié pour la commutation de signaux analogiques :

- résistance drain-source à l'état passant : $r < 100 \Omega$;
- résistance drain-source à l'état bloqué : $R > 20 M\Omega$;
- fréquence de commutation maximale : 1 MHz

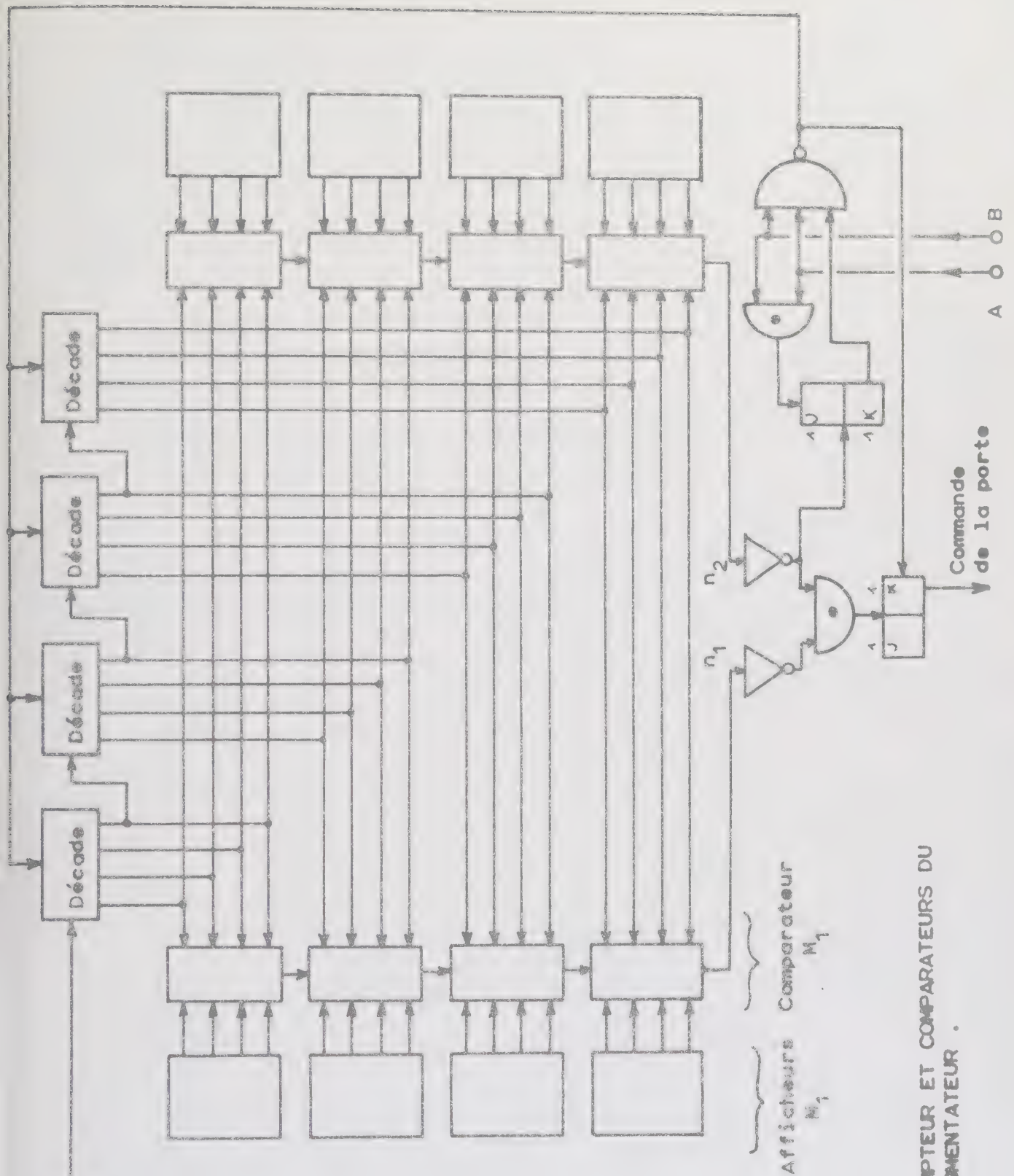


Figure 62 - COMPTEUR ET COMPAREURS DU SEGMENTATEUR .

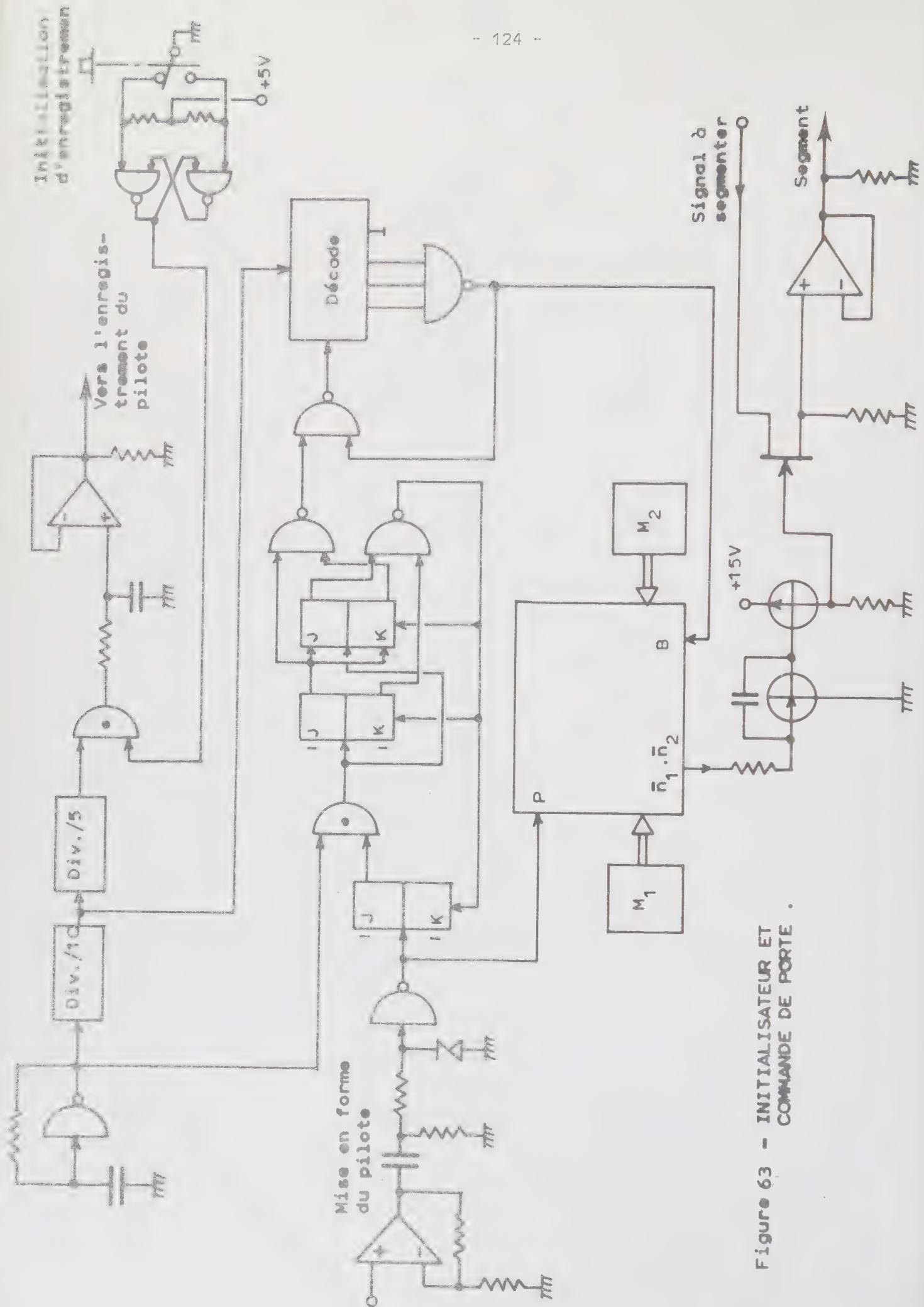


Figure 63 - INITIALISATEUR ET COMMANDE DE PORTE .

Le changement d'état du transistor est obtenu par la variation du potentiel (V_{GS}) de l'électrode "gate" :

- état passant : $V_{GS} = 0 \text{ V}$
- état bloqué : $V_{GS} = -15 \text{ V}$

Sa commande par la logique TTL utilisée nécessite donc un circuit d'interface permettant d'adapter les niveaux.

Les essais de cet élément nous ont montré de bonnes performances :

- bande passante : 0 à 60 kHz sans atténuation
- amplitude maximum du signal segmenté : 5 V.

Les figures 62 et 63 fournissent les schémas de réalisation de l'appareil.

L'amplificateur opérationnel de gain unité suivant la porte analogique, permet d'adapter l'impédance de sortie à 600 Ω . Son impédance d'entrée (R_E) doit être choisie de façon à ne pas affaiblir le signal lorsque la porte est passante, compte tenu de l'impédance de sortie de l'enregistreur et de la résistance du transistor, et à produire une atténuation d'au moins 60 dB lorsqu'elle est bloquée. Le choix de R_E d'environ 15 k Ω satisfait ces conditions.

La commutation des modes enregistrement-segmentation est interne, ce qui permet de manipuler sans avoir à modifier constamment le câblage du dispositif.

3 - CONCLUSION

L'utilisation du système de segmentation est simple :

- mode enregistrement : inscription simultanée sur les deux pistes de la boucle magnétique des impulsions du signal pilote et du signal à segmenter ; le bouton poussoir d'initialisation permet de définir le début d'enregistrement en inhibant le signal pilote.
- mode segmentation : l'enregistreur étant en position lecture, le segment apparaît automatiquement à la sortie de l'appareil ; il peut être synchronisé sur l'écran d'un oscilloscope en déclenchant le balayage de celui-ci par l'impulsion de fermeture de porte disponible à la sortie du comparateur correspondant à M_2 .

Pour obtenir la courbe de mélodie d'un segment de parole on enregistre le glottogramme sur la piste signal, et on branche la sortie du segmentateur à l'entrée de l'appareil de mesure de fréquence instantanée, dont la sortie est le résultat cherché.

Les instants d'ouverture et de fermeture de la porte peuvent être modifiés en cours de segmentation. Cet avantage, associé au tracé périodique automatique de la portion de courbe mélodique à chaque tour de boucle, permet de sélectionner avec précision l'élément voulu sur l'écran d'un oscilloscope, et de n'enregistrer graphiquement que l'information nécessaire.

C O N C L U S I O N

Les recherches relatives à la connaissance des langues et au fonctionnement de l'appareil phonatoire reposent, pour une large part, sur des analyses expérimentales des propriétés acoustiques du message parlé.

En particulier, la détermination précise en temps réel des variations de hauteur de la parole est considérée comme un problème important. Mais, jusqu'à ce jour, aucun dispositif alliant les performances techniques requises à une réalisation peu coûteuse et simple d'emploi n'a pu être proposé.

Notre étude, guidée par ces considérations, a permis de construire un nouveau type de "chaîne d'analyse de la mélodie de la parole", dont le principe est basé sur la mesure directe de la fréquence instantanée des émissions laryngées qui constituent la source des sons voisés. Le travail que nous avons résumé dans ce mémoire avait surtout pour but de confirmer la validité du principe et, au vu des performances du prototype, déduites de quelques expériences, de définir les domaines d'application intéressés qui sont, notamment :

- étude de l'intonation ;
- synthèse de mélodie ;
- analyse de la micromélodie ;
- étude des langues à tons ;
- apprentissage de schémas mélodiques.

Les tests que nous avons effectués sont très satisfaisants en ce qui concerne les caractéristiques fondamentales que doit présenter un "analyseur de mélodie" : fiabilité, précision, simplicité de conception et de mise en oeuvre.

Notre système ne peut cependant être considéré comme une réalisation définitive, et devra être modifié dans le sens d'une amélioration des circuits en fonction de l'évolution des techniques ; de plus, après une période d'expérimentation systématique sur des études concrètes, il sera souhaitable de l'adapter aux exigences des chercheurs.

En ce qui concerne plus spécialement le module de mesure logarithmique directe d'intervalles de temps successifs, l'étude de son principe original nous laisse entrevoir un développement intéressant, notamment dans son utilisation en "générateur" ; les domaines pour lesquels nous prévoyons une application sont, entre autres, l'acoustique musicale (génération de notes de référence) et l'instrumentation (génération de signaux à programmation dynamique logarithmique de la fréquence).

Le dispositif d'apprentissage de schémas mélodiques dont nous avons exposé le principe, devra également faire l'objet d'expériences suivies qui permettront d'en définir l'efficacité.

B I B L I O G R A P H I E

- [1] ABRAMSON (A.S)
The vowels and tones of standard Thai : acoustical measurements and experiments
I.J.A.L. 28(2), p.112-137, (1962)
- [2] ANDRE (P.) - LHOTE (F.) - FILLEAU (M)
Générateur digital de fréquences à affichage logarithmique
Congrès franco-allemand de chronométrie - Strasbourg (Mai 1975)
- [3] BILDSTEIN (P.)
Filtres actifs
Editions Radio (Paris, (1972)
- [4] BOË (L.J)
Introduction à la phonétique acoustique
Cours d'acoustique - Université de Grenoble (1970)
- [5] CARRE (R.)
Etude et réalisation d'un détecteur à fonctionnement automatique et continu des paramètres de la parole
Thèse (Faculté des Sciences de Grenoble), (1966)
- [6] CARRE (R.) - LANCIA (R.) - PAILLE (J.) - GSELL
Etude et réalisation d'un détecteur de mélodie pour analyse de la parole
L'onde Electrique, n° 434, p.556-562, (1963)

[7] CHEVILLARD (A.)

Contribution à l'étude analogique de l'appareil phonatoire

Thèse - Laboratoire d'Automatique de Besançon - (1973)

[8] DUVELLERDY (M.)

La glottographie en module et en phase, moyen d'étude des schémas électriques équivalents au larynx

Thèse de Médecine - Paris (1961)

[9] FABRE (P.)

Un procédé percutané d'inscription de l'accollement glottique au cours de la phonation : glottographie en H.F., premiers résultats

Bull. Acad. Nat. Méd., p.66-69, (1961)

[10] FABRE (P.)

Etude comparée des glottogrammes et des phonogrammes de la voix humaine

Ann. d'otorinolaryngologie, t.75, p.767-775, (1968)

[11] FABRE (P.)

La glottographie en H.F., particularités de l'appareillage

C.R. Soc. Biol., CL III, p.1361 (1959)

[12] FABRE (P.) - HUSSON (R.) - ROELENS (R.)

Etude expérimentale du comportement phonatoire des cordes vocales par la glottographie électrique

C.R. Acad. Sciences (Paris), t.248, p.3041 (1959)

[13] FANI (G.)

Acoustic theory of speech production

Mouton (The Hague), Paris, (1960)

[14] FARNSWORTH (D.W.)

High speed motion pictures of the human vocal cords

Bell Lab. Records, t.18, p.203-208, (1940)

- [15] FILLEAU (M.) - LHOTE (F.)
Mesure logarithmique de fréquence instantanée par un dispositif digital
 Revue d'Acoustique, n° 27, p.231-234, (1973)
- [16] FISCHER-JORGENSEN (E.) - FROKJAER-JENSEN (B.) - RISCHER (J.)
Preliminary experiments with the Fabre glottograph
 A.R.I.P.U.C. (Copenhagen), 1, p.22-31, (1966)
- [17] FLANAGAN (J.L.)
Speech analysis, synthesis and perception
 Berlin, Heidelberg, New-York, (1965)
- [18] FLANAGAN (J.L.)
Some properties of the glottal wave source
 J. Speech, Hearing Res., 1, p.99-116, (1958)
- [19] FLETCHER - WEGEL
The frequency sensitivity of normal ears
 Phys. Rev., 19, p.553, (1922)
- [20] FROKJAER-JENSEN (B.)
A photo electric glottograph
 A.R.I.P.U.C. (Copenhagen), 2, p.5-19, (1967)
- [21] FROKJAER-JENSEN (B.)
Comparison between a Fabre glottograph and a photo-electric glottograph
 A.R.I.P.U.C. (Copenhagen), 3, p.9-16, (1968)
- [22] GRÜTZMACHER (Von M.)
Demonstration eines Tonhöhenschreibers
 Phonetica 13, p.3-17, (1955)
- [23] HENRIOUD (J.M.) - LHOTE (F.)
Nouveau principe de conversion tension fréquence
 Annales Françaises de Chronométrie et de Micromécanique,
 p.27 à 33, (1972)

- [34] LIEBERMAN (P.)
Intonation, Perception and language
 Research Monograph, 38, M.I.T. Press, (1967)
- [35] MALCOLM
Une méthode d'extraction du fondamental
 L'Onde Electrique, p.110, (Mars 1973)
- [36] MALMBERG (B.)
Analyse des faits prosodiques - Problèmes et méthodes
 Cahiers de linguistique théorique et appliquée, Bucarest,
 p.123-127, 1966.
- [37] MARTINET (A.)
Eléments de linguistique générale
 Armand Colin, Collection U₂, 15, (1970)
- [38] METTAS (O.)
Etude sur l'intonation en français
 Trav. Ling. Litt. Strasbourg, 1, p.143-154, (1963)
- [39] MILLER (R.L.)
Nature of the vocal cord wave
 J.A.S.A., 31, p.667, (1959)
- [40] NOLL
Cepstrum pitch determination
 J.A.S.A., 41, p.293, (1967)
- [41] PAILLE (J.)
Contribution aux études sur la synthèse paramétrique de la parole. Synthétiseur à formants. Analogie de la source vocale.
 Thèse (Université Scientifique et Médicale de Grenoble), (1971)
- [42] PERSON (J.M.)
Détection et mesure de la fréquence fondamentale de la voix humaine par un algorithme câblé
 Revue d'Acoustique, n° 9, p.33 (1970)

- [43] PHILLIPS
Detection of the voice fundamental for visual pitch detection
J.A.S.A., 44, p.366, (1968)
- [44] RAISBECK
Voice pitch determination
J.A.S.A., 32, p.639, (1960)
- [45] ROUGET (G.)
Analyse des tons du gũ par le détecteur de mélodie
Revue langage et comportement, 1, p.37-47
- [46] SHOEDER (M.R.)
Period histogram and product spectrum : new methods for fundamental frequency measurement
J.A.S.A., 43, p.829, (1968)
- [47] SOBOLEV (V.N.)
Investigations expérimentales pour la détection automatique du ton fondamental par la méthode de corrélation
Journal d'Acoustique de l'Académie des Sciences d'U.R.S.S.
Vol.14, n° 13, (Janv.Mars 1969)
- [48] STRAKA
La division des sons en voyelles et consonnes peut-elle être justifiée ?
Trav. Ling. Univ. Strasbourg, 1, p.17-99, (1963)
- [49] TESTON (B.)
Description d'un système de détection de l'intensité et de la mélodie de la voix
Travaux de l'Institut de Phonétique d'Aix (Université de Provence), Volume 1, (1972)
- [50] THOMAS (J.M.C.)
Considérations sur le système des tons en Twi
La linguistique, 2, p.123-128, (1965)

- [51] THORVALDSEN (P.)
Construction of a segmentator
 A.R.I.P.U.C. (Copenhagen), 4, (1969)
- [52] TRAGER (G.L.)
Suprasegmental phonèmes of Thai (Bangkok dialect)
 M.A. Thesis (Georgetown), (1956)
- [53] TRAN HUONG MAI (A.)
Tones and intonation in South Vietnamese
 Ling. Circle of Cambera, Series A (9) South East Asian
 Linguistics (1), p.19-33, (1967)
- [54] VAISSIERE (J.)
Compte rendu de table ronde sur la synthèse par règles
 4èmes Journées d'Etude du Groupe de Communication parlée du
 G.A.L.F, Bruxelles, (1973)
- [55] VALLANCIEN (B.) - FAULHABER (J.)
What to think of glottography ?
 Folia Phoniatrica, 19, p.39-44, (1967)
- [56] VAN MICHEL (C.)
Les artefacts en glottographie
 Revue Electrodiagnostic, Thérapie, 4, (1967)
- [57] VAN MICHEL (C.) - RASKIN (L.)
L'électroglottomètre Mark 4. Son principe, ses possibilités
 Folia phoniatrica, 21, p.145-157, (1969)
- [58] VAN MICHEL (C.) - PFISTER (K.A.) - LUCHSINGER (R.)
Electroglottographie et Cinématographie laryngée ultra rapide
 Folia Phoniatrica, 22, p.81-91 (1970)
- [59] WAJSKOP (M.) - LANDERCY (A.) - SYLIN (G.)
*Etude et réalisation d'un segmentateur électronique et de son
 organe de commande*
 Revue d'acoustique n° 5, p.31-36, (1969)

[60] WEISS - HARRIS

Computer technique for high speed extraction of speech parameters
J.A.S.A., 35, p.207, (1963)

[61] ZIMA (P.)

On the function of tones in African languages
Travaux linguistiques de Prague, 2, p.151-156, (1966).

